United States
Department of
Agriculture

**Natural
Resources
Conservation
Service**

**Part 654 Stream Restoration Design
National Engineering Handbook**

Chapter 5 Stream Hydrology

Issued August 2007

Cover photo: Quantifying the flow of the stream involves analysis of rainfall/runoff, storm recurrence intervals, and watershed and flood plain conditions.

Advisory Note

Techniques and approaches contained in this handbook are not all-inclusive, nor universally applicable. Designing stream restorations requires appropriate training and experience, especially to identify conditions where various approaches, tools, and techniques are most applicable, as well as their limitations for design. Note also that product names are included only to show type and availability and do not constitute endorsement for their specific use.

Chapter 5 Stream Hydrology

Contents

Tables

Figures

Chapter 5 Stream Hydrology

654.0500 Purpose

Stream restoration design should consider a variety of flow conditions. These flows should be considered from both an ecological, as well as a physical perspective. Many sources and techniques for obtaining hydrologic data are available to the designer. This chapter provides a description of the flows and their analyses that should be considered for assessment and design. The computation of frequency distributions, with an emphasis on the log-Pearson distribution as provided in the U.S. Water Resources Council (WRC) Bulletin 17B, is addressed in detail. Examples have been provided to illustrate the methods. Transfer equations, risk, and low-flow methods are also addressed. Finally, this chapter describes advantages and limitations of four general approaches widely used for estimating the channel-forming discharge or dominant discharge.

654.0501 Introduction

Hydrologic analysis has historically been the starting point for channel design. Current and future flows were estimated, then the designer proceeded to further analysis. However, the complexities of stream restoration projects often require that hydrologic analysis be conducted in close coordination with a study of stream geomorphology and stream ecology.

Hydrologic computations are an integral part of any stream design and restoration project. However, design objectives for a stream restoration project cannot adequately be met by assessing channel behavior for only a single discharge. A stream restoration project usually has several design flows selected to meet various objectives. For example:

- Estimates of future flow conditions are often required to properly assess project performance over the long term.
- Estimates of low flows such as 7-day low flow often define critical habitat conditions.
- Estimates of channel-forming discharges are used to estimate stable channel dimensions.
- Flood flow estimates are used to assess stability of structures and flood plain requirements, as well as for scour depth prediction.

Many techniques are available to the designer for determining the various discharges used in assessment and design. The level of accuracy required for the different hydrologic analyses, as well as the need to estimate the different flows, is dependent on the site-specific characteristics of each project. Therefore, it is important to understand not only what each design flow represents, but also the underlying assumptions and the limitations of the techniques used to estimate the flow.

654.0502 Overview of design discharges

A description of some of the various types of design discharges is provided in this section. Although a project may not require the use of all of these flows for design, the hydraulic engineer/designer should still consider how the project will perform during a range of flow conditions.

(a) Low flows

Design of a low-flow channel may be required as part of a channel restoration. Normally, the design of the project for low flows is performed to meet biological goals. For instance, summer low flows are often a critical period for fish, and project goals may include narrowing the low-flow channel to provide increased depths at that time. Design flows may also be necessary to evaluate depths and velocities for fish spawning areas or fish passage during critical times of the year. Coordination with the biologist on the study team and familiarity with regulatory requirements are essential to make sure an appropriate flow (or range of flows) is selected.

(b) Channel-forming discharge

A determination of channel-forming discharge is used for many stability assessment tools and channel design techniques. The channel-forming discharge concept is based on the idea that for any given alluvial stream there exists a single discharge that, given enough time, would produce the width, depth, and slope equivalent to those produced by the natural flow in the stream. This discharge, therefore, dominates channel form and process. The channel-forming discharge concept evolved from the dominant discharge concept used to design irrigation canals in the latter part of the nineteenth and early part of the twentieth centuries. It is recognized, however, that the channel-forming discharge is a theoretical concept and may not be applicable to all stream types, especially flashy and ephemeral streams.

Depending on the application, channel-forming discharge can be estimated by several methods, based on:

- bankfull indices
- effective discharge
- specific recurrence interval
- drainage area

The distinction between channel-forming discharge and the other deterministic discharges is frequently confused, as the terms are used interchangeably. This chapter describes advantages and limitations of the four widely used general approaches.

(c) High discharge

The reaction of a channel to a high discharge can be the impetus for a stream restoration project. An identified high-flow event is often used in the design and specification of a design feature. The choice of a maximum design flow for stability analysis should be based on project objectives and consequences of failure. For example, the 100-year discharge might be used to design bank protection in a densely populated area, while a 10-year discharge might be appropriate in a rural stream. Other examples include:

- It may be a requirement to demonstrate that a proposed project will not raise the water surface profiles produced by a 5-year event (often referred to as nuisance-level flooding) sufficiently to adversely affect riparian infrastructure such as county roads, parks, and playgrounds.
- A significant flood event (typically no smaller than the 10-year frequency discharge) is used to estimate forces and compute scour depths at proposed habitat features constructed with logs. The goal is that these hard project features will withstand a flood of this magnitude without major damage, movement, or flanking.
- A significant flood event may have caused severe bank erosion, initiating a request to fix the erosion problems. It may be a requirement that any proposed fix provide stabilization that will be able to withstand a repetition of the forces produced by this event.
- It may be a requirement of the project design that the impacts of a 25-year flood event be limited to minor deposition of sediment and de-

bris; localized scour, erosion, and stone movement; and erosion of vegetation.

- Often the impact on the water surface profile for the 100-year flood event must be submitted as part of the project's permitting requirements. In many cases, it is a requirement to demonstrate that a proposed project will not result in increases to the 100-year flood plain area.
- It may also be necessary to estimate the flood-level reduction of a project on a 50-year flood event or for a larger event (such as the design discharge for a flood control project.).

(d) Flow duration

A flow-duration curve represents the percentage of time that a flow level is equaled or exceeded in a stream. This analysis is done for sediment transport assessments and ecological assessments, as well as for assessments of the duration of stress on soil bioengineering bank stabilization techniques.

Comparing flow-duration curves of different systems in a single basin or across a larger physiographic region can lend useful insight into a variety of watershed concerns. Issues such a flow contributions from ground water, watershed geology and geomorphology, and degree of flow regulation can also be examined, in part, with such a comparative analysis.

(e) Seasonal flows

It is often important to determine how the proposed restoration project will perform with low or normal flows. In addition, seasonal flow variations can have critical habitat importance. For example, a project goal may include a minimum flow depth during a critical spawning period for anadromous fish species and a lower minimum depth for resident fish species. The same techniques used to develop flow-duration curves for sediment analysis can also be used to assess and design for habitat conditions.

In many states, the U.S. Geological Survey (USGS) has developed regional regression curves for the critical flow periods. This might be the 10-year, 7-day low flow.

(f) Future flows

Estimates of future flow conditions are often required to properly assess future project performance. In some areas, the USGS has developed regional peak flow frequency curves that include a variable that can be used to estimate the impact of future changes in land use, such as an increase in the percent of impervious area for urban development. For example, typically 10 to 20 percent of the average rainfall event becomes runoff for an undeveloped watershed, while 60 to 70 percent of the average rainfall event becomes runoff for a developed (urbanized) watershed. However, regional equations typically do not include this variable, and a hydrologic model must be used to determine the change in the peak flow.

(g) Regulatory

Some Federal and state agencies have established minimum streamflow requirements for fish habitat. For example, the Federal Emergency Management Agency (FEMA) has established flood hazard maps for the 100-year and 500-year flood events and has estimated the flow associated with these events. Consultation with the appropriate authorities is needed if there is a possibility that a project will impact this flood level. Also, the U.S. Environmental Protection Agency (EPA) has established minimum flow requirements in many areas. These should be considered when determining the required design flows. While the determination and maintenance of these established flows may be based more on administrative decisions than current hydrologic data and analysis, they can be a critical component of a stream analysis or project design. A further description of regulatory requirements is provided in NEH654.17.

654.0503 Probability

Streamflow events are typically referred to by their return period. A return period of Rp means that in any given year, the event has a probability of occurrence (P):

$$P = \frac{1}{Rp} \qquad \text{(eq. 5–1)}$$

For example, a 100-year storm has an annual probability of occurrence of P=1/Rp=1/100=0.01 or 1 percent. Therefore, it is synonymous to speak of a 1 percent storm as a 100-year storm.

Risk is defined as the probability that one or more events will exceed a given magnitude within a specified period of years. Risk is calculated by means of the binomial distribution given in simplified form as follows:

$$R = 1 - (1 - P)^n \qquad \text{(eq. 5–2)}$$

where:
R = risk in decimal number
P = exceedance probability of event
n = number of years

The risk formula may be applied to many different scenarios, including the following:

- The likelihood of a 100-year flood occurring at least once in the next 100 years is 63 percent

$$R(100) = 1 - \left(1 - \frac{1}{100}\right)^{100} = 0.63$$

- The likelihood of a 100-year flood occurring at least once in the next 50 years is 39 percent

$$R(100) = 1 - \left(1 - \frac{1}{100}\right)^{50} = 0.39$$

- The likelihood of a 100-year event occurring at least once in 1,000 years is 99.996 percent, a very high probability, but never 100 percent.
- There is a 97 percent risk of a bankfull, 2-year recurrence interval discharge (50% annual chance) being exceeded in the next 5 years.
- Likewise, the 10-year discharge has a 41 percent risk of being exceeded in the next 5 years, or conversely, a 59 percent chance of not being exceeded.

Expected probability is a measure of the central tendency of the spread between confidence limits. Expected probability adjustment attempts to incorporate effects of uncertainty in application of frequency curves. The adjustment lessens as the stream record lengthens. Use of expected probability adjustment is often based on a policy decision.

It is important to note that a precipitation event may not have the same return period as a flow event. On small watersheds, a 100-year rainfall event may produce a 100-year flow or flood event. On large watersheds, however, the 100-year flow event may be produced by a series of smaller rainfall events. This distinction should particularly be kept in mind by the practitioner who is working with projects in large watersheds.

Equation 5–2 can also be rearranged to aid in determining a design storm (see example 1).

Example 1: Risk-based selection of design storm

Problem: A bank protection project involves considerable planting and soil bioengineering. However, the proposed planting would not be able to withstand the design storm until firmly established. The designer is asked to include reinforcement matting that will have a 90 percent chance of success over the next 5 years. What is the design storm?

Solution:

Step 1. Calculate the probability of an event occurring that is larger.

90 percent chance of success means that there is a 10 percent chance that an event will be larger.

Step 2. Rearrange equation 5–2 as follows:

$$R = 1 - (1 - P)^n$$

$$P = 1 - \sqrt[n]{(1 - R)}$$

And solve as:

$$P = 1 - \sqrt[5]{1 - 0.1}$$
$$= 0.0208$$

Step 3. Rearrange equation 5–1 as follows:

$$P = \frac{1}{R_p}$$

$$R_p = \frac{1}{P}$$

And solve as:

R_p = 47.9 or about a 50-year storm

654.0504 Gage analysis for flow frequency

Flow frequency analysis relates the magnitude of a given flow event with the frequency or probability of that event's exceedance. If a stream gage is available and the conditions applicable, a gage analysis is generally considered preferable, since it represents the actual rainfall-runoff behavior of the watershed in relation to the stream. A variety of Federal, state, and local agencies operate and maintain stream gages. Currently, the USGS operates about 7,000 active stream gaging stations across the country. Such data are also available for about 13,000 discontinued gaging stations. Historical peak flow data can be found at the following USGS Web site at:

http://nwis.waterdata.usgs.gov/usa/nwis/peak

It is important to determine if the present watershed conditions are represented by the stream gage record or if there has been a significant change in land use. If there has been a significant increase in urbanization, the historical record may not represent current conditions. While many hydrologic techniques are available for the prediction of frequency of flow events, this chapter presents concepts and techniques for analyzing peak flows and, to a lesser extent, low flows, following the recommendations of Guidelines for Determining Flood Flow Frequency, Bulletin 17B (WRC 1981).

Flow event data may be analyzed graphically or analytically. In graphical analysis, data are arrayed in order of magnitude, and each individual flow event is assigned a probability or recurrence interval (plotting point). The magnitudes of the flow events are then plotted against the probabilities, and a line or frequency curve is drawn to fit the plotted points. Peak flow data are usually plotted on logarithmic probability scales, which are spaced for the log-normal probability distribution to plot as a straight line. Data are often plotted to verify that the general trend agrees with frequency curves developed analytically.

(a) Analysis requirements and assumptions

In performing a frequency analysis of peak discharges, certain assumptions need to be verified including data independence, data sufficiency, climatic cycles and trends, watershed changes, mixed populations, and the reliability of flow estimates. The stream gage records must provide random, independent flow event data. These assumptions need to be kept in mind, otherwise, the resultant distribution of discharge frequencies may be significantly biased, leading to inappropriate designs and possible loss of property, habitat, and human life.

Data independence

To perform a valid peak discharge frequency analysis, the data points used in the analysis must be independent, that is, not related to each other. Flow events often occur over several days, weeks, or even months, such as for snowmelt. Only the peak discharge for each flow event should be used in the frequency analysis. Secondary peaks are dependent on each other and are not appropriate for use in a frequency analysis. Using secondary peaks would result in lower peak flows for a given frequency, since it would exaggerate the frequency of the magnitude of the event. It is common practice to minimize this problem by extracting annual peak flows from the streamflow record to use in the frequency analysis.

Data sufficiency

Gage records should contain a sufficient number of years of consecutive peak flow data. To minimize bias, this record should span both wet and dry years. In general, a minimum of 10 years is required (WRC 1981). However, longer gage records are generally recommended to estimate larger return periods and/or if there is a potential bias in the data set. This is addressed later in the climate bias example. If a gage record is shorter than optimum, it may be advisable to consider other methods of hydrologic estimations to support the gage analysis.

It is also important to use data that fully capture the peak for peak flow analysis. If a stream is flashy (typical of small watershed), the peak may occur over hours or even minutes, rather than days. If daily averages are used, then the flows may be artificially low and result in an underestimate of storm event values.

Therefore, for small watersheds, it may be necessary to look at hourly or even 15-minute peak data.

Annual duration
Gage analysis for flows with return intervals in excess of 2 years is typically conducted on annual series of data. This is the collection of the peak or maximum flow values that have occurred for each year in the duration of interest. Each year is defined by water year (Oct. 1 to Sept. 30).

Partial duration
When the desired event has a frequency of occurrence of less than 2 years, a partial duration series is recommended. This is a subset of the complete record where the analysis is conducted on values that are above a preselected base value. The base value is typically chosen so that there are no more than three events in a given year. In this manner, the magnitude of events that are equaled or exceeded three times a year can be estimated. Care must be taken to assure that multiple peaks are not associated with the same event, so that independence is preserved.

The return period for events estimated with the use of a partial duration series is typically 0.5 years less than what is estimated by an annual series (Linsley, Kohler, and Paulhus 1975). While this difference is fairly small for large events (100-yr for a partial vs. 100.5-yr for an annual series), it can be significant at more frequent events (1-yr for a partial vs. 1.5-yr for an annual series). Therefore, while an annual series may be sufficient to estimate the magnitude of a channel-forming discharge, it may not provide a precise estimate for the actual frequency of the discharge. It should also be noted that there is more subjectivity at the ends of both the annual and partial duration series frequency curves.

Climatic cycles and trends
Climatic cycles and trends have been identified in meteorological and hydrological records. Cycles in streamflow have been found in the world's major rivers. For example, Pekarova, Miklanek, and Pekar (2003) identified the following cycles of extreme river discharges throughout the world (years): 3.6, 7, 13, 14, 20 to 22, and 28 to 29. Some cycles have been associated with oceanic cycles, such as the El Niño-Southern Oscillation in the Pacific (Dettinger et al. 2000) and the North Atlantic Oscillation (Pekarova, Miklanek, and Pekar 2003). Trends in streamflow volumes and peaks are less apparent. However, trends in streamflow timing are likely, as have been presented in Cayan et al. (2001) for the western United States.

The identification of both cycles and trends is hampered by the relatively short records of streamflow available, as streamflow data increase, more cycles and trends may be identified. However, sufficient evidence does currently exist to warrant concern for the impact of climate cycles on the frequency analysis of peak flow data, even with 20, 30, or more years of record.

When performing a frequency analysis, it can be important to also analyze data at neighboring gages (that have longer or differing period of records) to assess the reasonableness of the streamflow data and frequency analysis at the site of interest. Keeping in mind the design life of the planned project and relating this to any climate cycles and trends identified during such a period, one can identify, in at least a qualitative manner, the appropriateness of the streamflow data. A case study is provided in example 2 that describes an analysis completed to assess climatic bias.

Paleoflood studies use geology, hydrology, and fluid dynamics to examine evidence often left by floods and may lead to a more comprehensive frequency analysis. Such studies are more relevant for projects with long design lives, such as dams. For more information on paleoflood techniques, see Ancient Floods, Modern Hazards: Principles and Applications of Paleoflood Hydrology (House et al. 2001).

Watershed changes
Land use and water use changes in watersheds can alter the frequency of high flows in streams. These changes, which are primarily caused by humans, include:

- urbanization
- reservoir construction, with the resulting attenuation and evaporation
- stream diversions
- construction of transportation corridors that increase drainage density
- deforestation from logging, infestation, high intensity fire
- reforestation

Example 2: Climatic bias case study

The Willow Creek watershed of the northern San Juan Mountains of Colorado has a wide range of stream-related projects being designed. This includes the remediation of drainage from tailings piles and mines; a braided to sinuous stream restoration; and the rehabilitation of a flume which carries flood flow through the town of Creede, Colorado. Discharge frequency estimates are necessary for all of these projects. The USGS had a gage operable in Creede for 32 years, from 1951 through 1982. Flow peaks measured for this 35.3-square-mile watershed ranged from 66 to 410 cubic feet per second. Thirty-two years of data is usually a reasonable record length for performing a frequency analysis. However, when six historic events were taken into account, the results of the 32-year frequency analysis appeared to be biased on the low side.

Records show that the historic events (with estimated peak flows of 1,200 ft^3/s and greater) occurred in the first half of the century in the Willow Creek watershed. This leads to a series of issues that should be examined:

- Were these historic peak estimates computed properly?
- Were these high flows random occurrences?
- Does this confliction indicate that all of the systematic record was recorded during a period of lesser precipitation and runoff?
- Or is some other mechanism occurring?

To shed more light on this situation an analysis was performed on two nearby, primarily undeveloped, watersheds: Carnero Creek, a 117-square-mile watershed with 78 years of record, and LaGarita Creek, a 61-square-mile watershed with 81 years of record. A sensitivity analysis was performed to assess the impact of a varying period of record. It was assumed that the records at these two locations cover three different periods: the actual period of record, the first half of the record, and the second half of the record. Frequency analyses were performed on each of these records. Results from this analysis are shown in table 5–1.

Table 5–1 Sensitivity analysis on gage record, Willow Creek case study

Gage ID	**08230500**			**Gage ID:**	**08231000**		
Stream	**Carnero Creek**			**Stream:**	**LaGarita Creek**		
Drainage area:	**117 mi^2**			**Drainage area:**	**61 mi^2**		
Years of record:	**1920–23, 26–28, 30, 32–2001**			**Years of record:**	**1920, 22–2001**		
	Log-Pearson results				**Log-Pearson results**		
	Full	**First half**	**Second half**		**Full**	**First half**	**Second half**
Record (yr)	78	39	39	Record (yr)	81	40	41
200-yr (ft^3/s)	1,690	1,780	695	200-yr (ft^3/s)	840	816	552
100-yr (ft^3/s)	1,290	1,470	554	100-yr (ft^3/s)	711	736	468
50-yr (ft^3/s)	958	1,180	435	50-yr (ft^3/s)	591	652	392
25-yr (ft^3/s)	694	921	333	25-yr (ft^3/s)	481	564	322
10-yr (ft^3/s)	424	618	223	10-yr (ft^3/s)	348	441	239
5-yr (ft^3/s)	271	417	155	5-yr (ft^3/s)	257	340	181
2-yr (ft^3/s)	118	187	80	2-yr (ft^3/s)	141	193	108
1.25-yr (ft^3/s)	53	79	43	1.25-yr (ft^3/s)	77	99	66

Example 2: Climatic bias case study—Continued

This analysis indicates sensitivity towards the specific period of record. Possible reasons for this bias include watershed changes such as forestry practices, climate cycles, and climate trends. For this Willow Creek example, additional stream gages within the region could be analyzed and extrapolated, using a regional regression methodology, to develop a more robust discharge frequency. A comparison of the computed discharges on a square-mile basis for selected discharges may show that the full record for the three stations is not that dissimilar.

For all frequencies, varying time periods used in a frequency analysis result in readily apparent differences. If a project's design were based on a frequency analysis for a gage with data gathered only during the second half of the twentieth century (as is the case in Willow Creek), this design may have attributes that are inappropriately sized.

Before a discharge-frequency analysis is used, or to judge how the frequency analysis is to be used, watershed history and records should be evaluated to assure that no significant watershed changes have occurred during the period of record. If such significant change has occurred, the period of record may need to be altered, or the frequency analysis may need to be used with caution, with full understanding of its limitations.

Particular attention should be paid to watershed changes when considering the use of data from discontinued gages. It was common to discontinue the small (<10 mi^2) drainage areas in the early 1980s. Aerial photographs can provide useful information in determining if the land use patterns of today are similar to those during the gage's period of record. Each gage site must be evaluated on an individual basis to determine whether today's watershed is still represented by yesteryear's flow records.

Mixed populations
At many locations, high flows are created by different types of events. For example, in mountain watersheds, high flow may result from snowmelt events, rain on snow events, or rain events. Also, tropical cyclones may produce differences from frontal systems. Gages with records that contain such different types of events require special treatment.

Reliability of flow estimates
Errors exist in streamflow records, as with all measured values. With respect to USGS records, data that are rated as excellent means that 95 percent of the daily discharges are within 5 percent of their true value, a good rating means that the data are within 10 percent of their true value, and a fair rating means that the data are within 15 percent of their true value. Records with greater than 15 percent error are considered poor (USGS 2002b).

These gage inaccuracies are often random, possibly minimizing the resultant error in the frequency analysis. Overestimates may be greatest for larger, infrequent events, especially the historic events. For example, research indicates that mobile bed streams cannot maintain supercritical flow over long distances and time periods. Therefore, a critical flow assumption is more appropriate in these situations. For more information on these methods, see Grant (1997) and Webb and Jarrett (2002). If consistent overestimation has occurred, the error is not random, but is instead, a systematic bias.

Regulated flows
Flows from dams are considered to be regulated flows. The normal statistical techniques in Flood Flow Frequency, Bulletin No. 17B (WRC 1981) cannot be used in these situations. However, in some cases, standard graphic statistical techniques can be used to determine the frequency curve. A review of the reservoir operation plan and project design document will provide information on the downstream releases.

(b) Frequency distributions

A flow frequency analysis is a consistent, statistical method for denoting the probability of occurrence of flows at a specific point in a stream system. Such relationships are required in the planning and design of structures in and near streams. However, peak flow frequency analysis techniques have limitations as described in NEH654.0502. Until hydrologic process modeling becomes more developed, the use of the following statistical methods is necessary.

Statistical parameters
The basic statistical parameters used in frequency analyses (applied to both normal and logarithmic values) are:

Mean:

$$\text{Mean: } \bar{X} = \frac{\sum X}{n} \qquad \text{(eq. 5–3)}$$

Standard deviation:

$$\text{Standard deviation: } S = \left[\frac{\sum\left(X - \bar{X}\right)^2}{n-1}\right]^{0.5} \qquad \text{(eq. 5–4)}$$

Skew coefficient:

$$\text{Skew coefficient: } G = \frac{n\sum\left(X - \bar{X}\right)^3}{(n-1)(n-2)S^3} \qquad \text{(eq. 5–5)}$$

where:
X = annual peak flow or logarithm of annual peak flow
n = length of data set

The mean is the arithmetic average of the data. It is the expected value of the data. The standard deviation is essentially an indication of how much the data is spread about the mean. The smaller the standard deviation value, the closer are the data points to the mean. For a normally distributed data set, approximately two-thirds of the data will be within plus or minus one standard deviation of the mean, while almost 95 percent will be within two standard deviations of the mean. Skewness is the third central moment about the mean and a measure of symmetry (or rather the lack of symmetry) of a data set (Fripp, Fripp, and Fripp 2003). If values are further from the mean on one side than the other, the distribution will have a larger skew. The skew has a large effect on the shape, and thus, the value of a distribution. Transformations (such as converting to logarithmic forms) are often made on skewed data. Spreadsheets are commonly used to compute these parameters.

Common distributions

Four distributions are most common in frequency analyses of hydrologic data, specifically the normal distribution, log-normal distribution, Gumbel extreme value distribution, and log-Pearson type III distribution. The log-Pearson distribution has been recommended by the WRC and is the primary method for discharge-frequency analyses in the United States. It is also recommended in NEH630.18.

However, the use of the log-Pearson distribution is not universal. For example, Great Britain and China use the generalized extreme value distribution and the log-normal distribution, respectively, while other countries commonly use other distributions (Singh and Strupczewski 2002). This section presents an overview of the four distributions. However, only the log-Pearson distribution will be addressed in detail.

Normal distribution

The normal or Gaussian distribution is one of the most popular distributions in statistics. It is also the basis for the log-normal distribution, which is often used in hydrologic applications. The distribution, as used in frequency analysis computations, is provided:

$$X_{N,T} = \overline{X} + K_{N,T}S \qquad \text{(eq. 5–6)}$$

where:

$X_{N,T}$ = predicted discharge, at return period T

$\overline{X}$ = average annual peak discharge

$K_{N,T}$ = normal deviate (z) for the standard normal curve, where area = $0.50 - \frac{1}{T}$

S = standard deviation, of annual peak discharge

Log-normal distribution

The annual maximum flow series is usually not well approximated by the normal distribution; it is skewed to the right, since flows are only positive in magnitude, while the normal distribution includes negative values. When a data series is left-bounded and positively skewed, a logarithmic transformation of the data may allow the use of normal distribution concepts through the use of the log-normal distribution. This transformation can correct this problem through the conversion of all flow values to logarithms. This is the method used in the log-normal distribution:

$$X_{LN,T} = \overline{X}_l + K_{LN,T}S_l \qquad \text{(eq. 5–7)}$$

where:

$X_{LN,T}$ = logarithm of predicted discharge, at return period

$\overline{X}_l$ = average of annual peak discharge logarithms

$K_{LN,T}$ = normal deviate (z), of logarithms for the standard normal curve, where area = $0.50 - \frac{1}{T}$

S_l = standard deviation, of logarithms of annual peak discharge

Gumbel extreme value distribution

Peak discharges commonly have a positive skew, because one or more high values in the record result in the distribution not being log-normally distributed. Hence, the Gumbel extreme value distribution was developed.

$$X_{G,T} = \overline{X}_l + K_{G,T}S \qquad \text{(eq. 5–8)}$$

where:

$X_{G,T}$ = predicted discharge, at return period T

$\overline{X}_l$ = average annual peak discharge

$K_{G,T}$ = a function of return period and sample size, provided in table 5–2

S = standard deviation of annual peak discharge

Log-Pearson Type III distribution

The log-Pearson type III distribution applies to nearly all series of natural floods and is the most commonly used frequency distribution for peak flows in the United States. It is similar to the normal distribution, except that the log-Pearson distribution accounts for the skew, instead of the two parameters, standard deviation and mean. When the skew is small, the log-Pearson distribution approximates a normal distribution. The basic distribution is:

$$X_{LP,T} = \overline{X}_l + K_{LP,T} S_l \quad \text{(eq. 5–9)}$$

where:

$X_{LP,T}$ = logarithm of predicted discharge, at return period T

$\overline{X}$ = average of annual peak discharge logarithms

$K_{LP,T}$ = a function of return period and skew coefficient, provided in table 5–3

S_l = standard deviation of logarithms of annual peak discharge

The mean in a log-Pearson type III distribution is approximately equal to the logarithm of the 2-year peak discharge. The standard deviation is the slope of the line, and the skew is shown by the curvature of the line.

The log-Pearson type III distribution has been recommended by the WRC, and the NRCS has adopted its use (NEH630.18). Details on the use of the log-Pearson distribution for the determination of flood frequency are presented later in this chapter. More information is also available in WRC Bulletin 17B.

Plotting position

The graphical evaluation of the adequacy-of-fit of a frequency distribution is recommended when performing an analysis. Plotting positions are used to estimate the return period of actual annual peak flows in these plots. The Weibull equation is provided:

Weibull: $$PP = \frac{100m}{n} \quad \text{(eq. 5–10)}$$

where:

n = sample size

m = data rank

The basic computations for a discharge-frequency analysis are illustrated in example 3.

Application of log-Pearson frequency distribution

Numerous statistical distributions that can provide a fit of annual peak flow data exist. The hydrology committee of the WRC (WRC 1981) recommended the use of the log-Pearson type III distribution because it provided the most consistent fit of peak flow data. NRCS participated on the hydrology committee and has adopted the use of WRC Bulletin 17B for determining flood flow frequency, using measured streamflow data.

Several computer programs and Microsoft® Excel® spreadsheet programs exist that can be used to perform log-Pearson frequency analysis. A spreadsheet example is used in many of the examples in this chapter.

Table 5–2 K-values for the Gumbel extreme value distribution

Sample size	Return period, T (yr)								
	1.11	**1.25**	**2.00**	**2.33**	**5**	**10**	**25**	**50**	**100**
15	–1.34	–0.98	–0.15	0.06	0.97	1.70	2.63	3.32	4.01
20	–1.29	–0.95	–0.15	0.05	0.91	1.63	2.52	3.18	3.84
25	–1.26	–0.93	–0.15	0.04	0.89	1.58	2.44	3.09	3.73
30	–1.24	–0.91	–0.16	0.04	0.87	1.54	2.39	3.03	3.65
40	–1.21	–0.90	–0.16	0.03	0.84	1.50	2.33	2.94	3.55
50	–1.20	–0.88	–0.16	0.03	0.82	1.47	2.28	2.89	3.49
60	–1.18	–0.87	–0.16	0.02	0.81	1.45	2.25	2.85	3.45
70	–1.17	–0.87	–0.16	0.02	0.80	1.43	2.23	2.82	3.41
80	–1.16	–0.86	–0.16	0.02	0.79	1.42	2.21	2.80	3.39
100	–1.15	–0.85	–0.16	0.02	0.77	1.40	2.19	2.77	3.35
200	–1.11	–0.82	–0.16	0.01	0.74	1.33	2.08	2.63	3.18
400	–1.07	–0.80	–0.16	0.00	0.70	1.27	1.99	2.52	3.05

Table 5–3 K-values for the log-Pearson type III distribution

Skew coefficient C_S	Recurrence interval/percent chance of occurrence								
	1.0526 **95**	**1.25** **80**	**2** **50**	**5** **20**	**10** **10**	**25** **4**	**50** **2**	**100** **1**	**200** **0.5**
–2.00	–1.996	–0.609	0.307	0.777	0.895	0.959	0.980	0.990	0.995
–1.90	–1.989	–0.627	0.294	0.788	0.920	0.996	1.023	1.307	1.044
–1.80	–1.981	–0.643	0.282	0.799	0.945	1.035	1.069	1.087	1.097
–1.70	–1.972	–0.660	0.268	0.808	0.970	1.075	1.116	1.140	1.155
–1.60	–1.962	–0.675	0.254	0.817	0.994	1.116	1.166	1.197	1.216
–1.50	–1.951	–0.690	0.240	0.825	1.018	1.157	1.217	1.256	1.282
–1.40	–1.938	–0.705	0.225	0.832	1.041	1.198	1.270	1.318	1.351
–1.30	–1.925	–0.719	0.210	0.838	1.064	1.240	1.324	1.383	1.424
–1.20	–1.910	–0.732	0.195	0.844	1.086	1.282	1.379	1.449	1.501
–1.10	–1.894	–0.745	0.180	0.848	1.107	1.324	1.435	1.518	1.581
–1.00	–1.877	–0.758	0.164	0.852	1.128	1.366	1.492	1.588	1.664
–0.90	–1.858	–0.769	0.148	0.854	1.147	1.407	1.549	1.660	1.749
–0.80	–1.839	–0.780	0.132	0.856	1.166	1.448	1.606	1.733	1.837
–0.70	–1.819	–0.790	0.116	0.857	1.183	1.488	1.663	1.806	1.926
–0.60	–1.797	–0.800	0.099	0.857	1.200	1.528	1.720	1.880	2.016
–0.50	–1.774	–0.808	0.083	0.856	1.216	1.567	1.777	1.955	2.108
–0.40	–1.750	–0.816	0.066	0.855	1.231	1.606	1.834	2.029	2.201
–0.30	–1.726	–0.824	0.050	0.853	1.245	1.643	1.890	2.104	2.294
–0.20	–1.700	–0.830	0.033	0.850	1.258	1.680	1.945	2.178	2.388
–0.10	–1.673	–0.836	0.017	0.846	1.270	1.716	2.000	2.252	2.482
0.00	–1.645	–0.842	0.000	0.842	1.282	1.751	2.054	2.326	2.576
0.10	–1.616	–0.846	–0.017	0.836	1.292	1.785	2.107	2.400	2.670
0.20	–1.586	–0.850	–0.033	0.830	1.301	1.818	2.159	2.472	2.763
0.30	–1.555	–0.853	–0.050	0.824	1.309	1.849	2.211	2.544	2.856
0.40	–1.524	–0.855	–0.066	0.816	1.317	1.880	2.261	2.615	2.949
0.50	–1.491	–0.856	–0.083	0.808	1.323	1.910	2.311	2.686	3.041
0.60	–1.458	–0.857	–0.099	0.800	1.328	1.939	2.359	2.755	3.132

Table 5–3 K-values for the log-Pearson type III distribution—Continued

Skew coefficient C_S	Recurrence interval/percent chance of occurrence								
	1.0526	**1.25**	**2**	**5**	**10**	**25**	**50**	**100**	**200**
	95	**80**	**50**	**20**	**10**	**4**	**2**	**1**	**0.5**
0.70	–1.423	–0.857	–0.116	0.790	1.333	1.967	2.407	2.824	3.223
0.80	–1.388	–0.856	–0.132	0.780	1.336	1.993	2.453	2.891	3.312
0.90	–1.353	–0.854	–0.148	0.769	1.339	2.018	2.498	2.957	3.401
1.00	–1.317	–0.852	–0.164	0.758	1.340	2.043	2.542	3.022	3.489
1.10	–1.280	–0.848	–0.180	0.745	1.341	2.066	2.585	3.087	3.575
1.20	–1.243	–0.844	–0.195	0.732	1.340	2.087	2.626	3.149	3.661
1.30	–1.206	–0.838	–0.210	0.719	1.339	2.108	2.666	3.211	3.745
1.40	–1.168	–0.832	–0.225	0.705	1.337	2.128	2.706	3.271	3.828
1.50	–1.131	–0.825	–0.240	0.690	1.333	2.146	2.743	3.330	3.910
1.60	–1.093	–0.817	–0.254	0.675	1.329	2.163	2.780	3.388	3.990
1.70	–1.056	–0.808	–0.268	0.660	1.324	2.179	2.815	3.444	4.069
1.80	–1.020	–0.799	–0.282	0.643	1.318	2.193	2.848	3.499	4.147
1.90	–0.984	–0.788	–0.294	0.627	1.310	2.207	2.881	3.553	4.223
2.00	–0.949	–0.777	–0.307	0.609	1.302	2.219	2.912	3.605	4.398
2.10	–0.914	–0.765	–0.319	0.592	1.294	2.230	2.942	3.656	4.372
2.20	–0.882	–0.752	–0.330	0.574	1.284	2.240	2.970	3.705	4.444
2.30	–0.850	–0.739	–0.341	0.555	1.274	2.248	2.997	3.753	4.515
2.40	–0.819	–0.725	–0.351	0.537	1.262	2.256	3.023	3.800	4.584
2.50	–0.790	–0.711	–0.360	0.518	1.250	2.262	3.048	3.845	4.652
2.60	–0.762	–0.696	–0.368	0.499	1.238	2.267	3.071	3.889	4.718
2.70	–0.736	–0.681	–0.376	0.479	1.224	2.272	3.093	3.932	4.783
2.80	–0.711	–0.666	–0.384	0.460	1.210	2.275	3.114	3.973	4.847
2.90	–0.688	–0.651	–0.390	0.440	1.195	2.277	3.134	4.013	4.909
3.00	–0.665	–0.636	–0.396	0.420	1.180	2.278	3.152	4.051	4.970

Example 3: Example computations of a discharge-frequency analysis

Problem: A streambank stabilization project is being designed for the Los Pinos River in the San Luis Valley of the Upper Rio Grande Basin. The project is less than a mile downstream from a USGS stream gage. The 167-square-mile watershed consists of forests, grass, and sage on a rural, primarily public land setting. This problem illustrates the analysis of the gaged flow data with the four common distributions.

Solution: First, the gage information is downloaded from the USGS (*http://waterdata.usgs.gov/nwis*), and the data are sorted and transformed. Then the basic statistics are calculated (table 5–4).

After computing these basic statistics, the distributions can be generated. Specifically, the magnitudes of the 1.25-, 2-, 5-, 10-, 25-, 50-, and 100-year events are computed and plotted with the source data. Example computation and plotting of distributions are shown in figure 5–1. Both the Gumbel and log-Pearson distributions fit the plotted data reasonably well.

Table 5–4 Discharge peaks, with basic statistics

USGS 08248000 Los Pinos River near Ortiz, CO

Year	Peak discharge (ft^3/s)	ln (peak discharge)	Rank	Weibull plotting position (yr)
1915	1,620	7.3902	27	3.1
1916	1,690	7.4325	22	3.8
1917	1,750	7.4674	18	4.7
1918	1,020	6.9276	57	1.5
1919	1,550	7.3460	30	2.8
1920	2,300	7.7407	5	16.8
1925	1,160	7.0562	50	1.7
1926	1,600	7.3778	29	2.9
1927	1,680	7.4265	23	3.7
1928	1,240	7.1229	47	1.8
1929	1,180	7.0733	48	1.8
1930	1,100	7.0031	51	1.6
1931	684	6.5280	72	1.2
1932	2,000	7.6009	10	8.4
1933	1,490	7.3065	35	2.4
1934	569	6.3439	75	1.1
1935	1,420	7.2584	40	2.1
1936	1,640	7.4025	26	3.2
1937	2,770	7.9266	4	21.0
1938	2,270	7.7275	6	14.0
1939	1,360	7.2152	43	2.0
1940	887	6.7878	63	1.3
1941	3,160	8.0583	1	84.0

Year	Peak discharge (ft^3/s)	ln (peak discharge)	Rank	Weibull plotting position (yr)
1942	2,000	7.6009	10	8.4
1943	1,370	7.2226	42	2.0
1944	3,030	8.0163	2	42.0
1945	2,180	7.6871	7	12.0
1946	1,090	6.9939	54	1.6
1947	1,740	7.4616	19	4.4
1948	1,660	7.4146	24	3.5
1949	1,620	7.3902	27	3.1
1950	876	6.7754	65	1.3
1951	563	6.3333	77	1.1
1952	2,790	7.9338	3	28.0
1953	924	6.8287	62	1.4
1954	882	6.7822	64	1.3
1955	700	6.5511	71	1.2
1956	926	6.8309	61	1.4
1957	1,850	7.5229	14	6.0
1958	1,490	7.3065	35	2.4
1959	646	6.4708	73	1.2
1960	1,100	7.0031	51	1.6
1961	1,420	7.2584	40	2.1
1962	1,480	7.2998	37	2.3
1963	532	6.2766	78	1.1
1964	1,000	6.9078	60	1.4

Table 5–4 Discharge peaks, with basic statistics—Continued

Year	Peak discharge (ft^3/s)	ln [peak discharge]	Rank	Weibull Plotting Position (yr)
1965	2,000	7.6009	10	8.4
1966	1,010	6.9177	59	1.4
1967	755	6.6267	70	1.2
1968	1,340	7.2004	44	1.9
1969	1,180	7.0733	48	1.8
1970	1,500	7.3132	34	2.5
1971	488	6.1903	81	1.0
1972	385	5.9532	82	1.0
1973	1,940	7.5704	13	6.5
1974	841	6.7346	68	1.2
1975	2,020	7.6109	8	10.5
1976	1,060	6.9660	55	1.5
1977	379	5.9375	83	1.0
1978	1,050	6.9565	56	1.5
1979	1,810	7.5011	15	5.6
1980	1,660	7.4146	24	3.5
1981	580	6.3630	74	1.1
1982	1,530	7.3330	32	2.6
1983	1,700	7.4384	21	4.0
1984	1,790	7.4900	16	5.3
1985	2,020	7.6109	8	10.5
1986	1,710	7.4442	20	4.2
1987	1,430	7.2654	39	2.2
1988	501	6.2166	80	1.1
1989	860	6.7569	66	1.3
1990	564	6.3351	76	1.1
1991	1,470	7.2930	38	2.2
1992	845	6.7393	67	1.3
1993	1,780	7.4844	17	4.9
1994	1,540	7.3395	31	2.7
1995	1,510	7.3199	33	2.5
1996	840	6.7334	69	1.2
1997	1,340	7.2004	44	1.9
1998	1,100	7.0031	51	1.6
1999	1,020	6.9276	57	1.5
2000	516	6.2461	79	1.1
2001	1,300	7.1701	46	1.8

For Q: average = 1,366
standard deviation = 598
skew coefficient = 0.660

For lnQ: average = 7.1165
standard deviation = 0.4763
skew coefficient = -0.507

For lnQ: average = 7.1165
standard deviation = 0.4763
skew coefficient = -0.507

Figure 5–1 Plotting distributions for return period peak discharges

For the normal distribution, equation 5–6 is used to compute the peaks. But first, a table of areas under the standard normal curve (found in most statistics books) is used to determine the K_N and K_{LN} values.

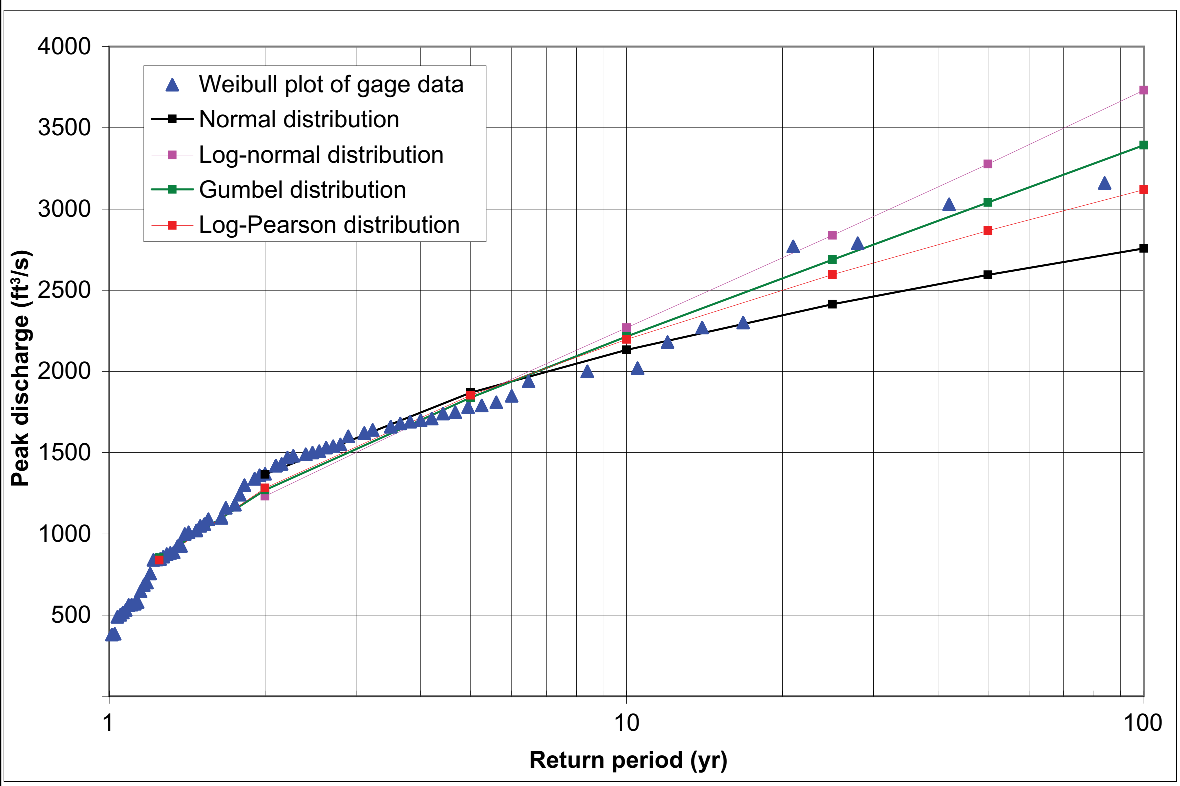
Using the equation $K_{N,T} = K_{LN,T}$ = 0.50¢–1/T, the following table is populated:

Return Period	1.25	2.00	5.00	10.00	25.00	50.00	100.00
K_N & K_{LN}	----	0.00	0.84	1.28	1.75	2.05	2.33

With these K-values and addiing the K-values for the Gumbel and Log-Pearson distributions, the following table is generated, using equations 5–6 through 5–7.

Q: average:	1366	**ln Q:** ıverage:	7.12
standard deviation:	598	standard deviation:	0.48
skew coefficient:	0.66	skew coefficient:	-0.51

Method		Return Period						
		1.25	2	5	10	25	50	100
Normal Distribution	K_N	----	0.00	0.84	1.28	1.75	2.05	2.33
	Q_N (ft³/s)	----	1,366	1,869	2,132	2,413	2,594	2,757
Log-normal distribution	K_{LN}	----	0.00	0.84	1.28	1.75	2.05	2.33
	Q_{LN}	----	7.12	7.52	7.73	7.95	8.09	8.22
	Q_{LN} (ft³/s)	----	1,232	1,840	2,269	2,837	3,277	3,732
Gumbel distribution	K_G	-0.86	-0.16	0.79	1.42	2.21	2.80	3.39
	Q_G (ft³/s)	852	1,270	1,838	2,215	2,688	3,040	3,393
Log-Pearson distribution	K_{LP}	-0.81	0.08	0.86	1.21	1.56	1.77	1.95
	Q_{LP}	6.73	7.16	7.52	7.70	7.86	7.96	8.05
	Q_{LP} (ft³/s)	839	1,282	1,852	2,198	2,596	2,867	3,119



General log-Pearson distribution

This distribution was provided previously as equation 5–9. The average, standard deviation, and skew coefficient were defined by equations 5–5 through 5–7. A complete table of K-values, with skews from –9.0 to +9.0, can be obtained from appendix 3 of WRC Bulletin 17B.

Generalized skew and weighting the skew coefficient

The computed station skew is sensitive to large events, especially with short periods of records. This problem can be minimized by weighting the station skew with a generalized skew that takes into account skews from neighboring gaged watersheds.

Three methods to develop this generalized skew are to:

- develop a skew isoline map
- develop a skew regression (or prediction) equation
- compute the mean and variance of the skew coefficients

These methods should incorporate at least 40 stations with at least 25 years of record within the gage of interest's hydro-physiographic province. Plate 1 of WRC Bulletin 17B could also be used; but, due to the vintage of this compilation, a detailed study may be preferred.

To develop a skew isoline map, the station skews are plotted at the centroid of the watershed and trends are observed. A regression or prediction equation can also be developed to relate skews to watershed and climatologic characteristics. If no relationship can be found with the isoline or regression approach, the arithmetic mean ($\overline{X}$) can be simply computed and used as the generalized skew. Care needs to be taken to ensure that all of the gages are in a similar hydro-physiographic province.

Once the best generalized skew is computed, a weighted skew is computed for the log-Pearson analysis using equation. 5–11.

$$G_W = \frac{S^2_{\overline{G}}\left(G\right)+S^2_G\left(\overline{G}\right)}{S^2_{\overline{G}}+S^2_G} \qquad \text{(eq. 5–11)}$$

where:
G_W = weighted skew coefficient
G = station skew
$\overline{G}$ = generalized skew
$S^2_{\overline{G}}$ = variance (mean square error) of generalized skew
S^2_G = variance of station skew

When generalized skews are read from Plate 1 of WRC Bulletin 17B, a variance of 0.302 should be used in equation 5–11.

The variance of the logarithmic station skew is a function of record length and population skew. This variance can be approximated with equations 5–12, 5–13, and 5–14.

$$S^2_G = 10^{\left[A-B\left(\log_{10}\left(\frac{n}{10}\right)\right)\right]} \qquad \text{(eq. 5–12)}$$

$$\begin{aligned} A &= -0.33+0.08\left|G\right| && \text{if } \left|G\right|\le 0.90 \\ A &= -0.52+0.30\left|G\right| && \text{if } \left|G\right|> 0.90 \qquad \text{(eq. 5–13)} \\ B &= 0.94-0.26\left|G\right| && \text{if } \left|G\right|\le 1.50 \end{aligned}$$

$$B = 0.55 \qquad \text{if } \left|G\right|>1.50 \qquad \text{(eq. 5–14)}$$

where:
n = record length in years
$\left|G\right|$ = absolute value of the station skew

If an historic record adjustment has been made. Historically adjusted values should be used.

Broken or incomplete records

Annual peaks for certain years at a gage are often missing. If this happens, the two or more record lengths are analyzed as a continuous record, with a record length equal to the sum of individual records.

Incomplete records refer to a high or low streamflow record that is missing due to a gaging failure. Usually, the gaging agency uses an indirect flow estimate to fill this void. If this has not occurred, effort to fill this gap may be warranted.

Historic flood data

As described in reliability of flow estimates, high flow values and historic events can be overestimated. If historic data are judged not to be biased high, WRC Bulletin 17B provides a special procedure for dealing with these events, instead of using a broken record approach. This method assumes that data from the systematic record is representative of the period between the historic data, and the systematic record and its statistics are adjusted accordingly.

First, a systematic record weight is computed.

$$W_s = \frac{H - Z}{n + L} \qquad \text{(eq. 5–15)}$$

where:
W_s = systematic record weight
H = historical period
Z = number of historic peaks
n = systematic record length
L = number of low values excluded, including low outliers and zero flow years

The historically adjusted average ($\tilde{X}$) is computed using:

$$\tilde{X} = \frac{W_s \sum X_s + \sum X_h}{H - W_s L} \qquad \text{(eq. 5–16)}$$

where:
X_s = logarithmic systematic record peaks
X_h = logarithmic historic record peaks

The historically adjusted standard deviation of logarithms ($\tilde{S}$) is:

$$\tilde{S} = \sqrt{\tilde{S}^2} = \sqrt{\frac{W_s \sum \left(X - \tilde{X}\right)^2 + \sum \left(X_h - \tilde{X}\right)^2}{\left(H - W_s L - 1\right)}} \qquad \text{(eq. 5–17)}$$

The historically adjusted skew coefficient of logarithms ($\tilde{G}$) is:

$$\tilde{G} = \frac{H - W_s L - 1}{\left(H - W_s L - 1\right)\left(H - W_s L - 2\right)} \left[\frac{W_s \sum \left(X - \tilde{X}\right)^3 + \sum \left(X_h - \tilde{X}\right)^3}{\tilde{S}^3}\right] \qquad \text{(eq. 5–18)}$$

Outliers

Outliers are data points that depart significantly from the trend of the remaining data. Including such outliers may be inappropriate in a frequency analysis. The decision to retain or eliminate an outlier is based on both hydrologic and statistical considerations. The statistical method for identifying possible outliers, as presented in WRC Bulletin 17B, uses equations 5–19 and 5–20:

$$X_H = \bar{X}_l + K_o S_l \qquad \text{(eq. 5–19)}$$

$$X_L = \bar{X}_l - K_o S_l \qquad \text{(eq. 5–20)}$$

where:
X_H = high outlier threshold, in logarithm units
X_L = low outlier threshold, in logarithm units (no historic adjustment)
$\bar{X}_l$ = average of annual peak discharge logarithms
K_o = based on sample size n, as listed in table 5–5
S_l = standard deviation of logarithms of annual peak discharge

If the station skew is greater than +0.4, high outliers are considered first and possibly eliminated. If the station skew is less than –0.4, low outliers are considered first, and then possibly eliminated. When the skew is between ± 0.4, a test for both high and low outliers should be first applied before possibly eliminating any outliers from the data set.

If an adjustment for historic flood data has already been made, the low outlier threshold equation is modified in the form:

$$X_{L,H} = \tilde{X} - K_o \tilde{S} \qquad \text{(eq. 5–21)}$$

where:
$X_{L,H}$ = low outlier threshold, in logarithm units (with historic adjustment)
$\tilde{X}$ = historically adjusted mean logarithm
$\tilde{S}$ = historically adjusted standard deviation

Mixed populations

In many watersheds, annual peak flows are caused by different types of events such as snowmelt, tropical cyclones, and summer thunderstorms. Including all types of events in a single frequency analysis may result in large and inappropriate skew coefficients. For such situations, special treatment may be warranted. Specifically, peak flows can be segregated by cause, analyzed separately, and then combined. Importantly, separation by calendar date alone is not appropriate, unless it can be well documented that an event type always varies by time of year.

Zero flow years

Some streams in arid regions may have no flow during the entire water year, thus having one or more zero peak flow values in its record. Such situations require special treatment. See appendix 5 in WRC Bulletin 17B

Table 5–5 Outlier test K_o values, from WRC Bulletin 17B

Sample size	K_o value	Sample size	K_o value	Sample size	K_o value	Sample size	K_o value
10	2.036	45	2.727	80	2.940	115	3.064
11	2.088	46	2.736	81	2.945	116	3.067
12	2.134	47	2.744	82	2.949	117	3.070
13	2.175	48	2.753	83	2.953	118	3.073
14	2.213	49	2.760	84	2.957	119	3.075
15	2.247	50	2.768	85	2.961	120	3.078
16	2.279	51	2.775	86	2.966	121	3.081
17	2.309	52	2.783	87	2.970	122	3.083
18	2.335	53	2.790	88	2.973	123	3.086
19	2.361	54	2.798	89	2.977	124	3.089
20	2.385	55	2.804	90	2.981	125	3.092
21	2.408	56	2.811	91	2.984	126	3.095
22	2.429	57	2.818	92	2.989	127	3.097
23	2.448	58	2.824	93	2.993	128	3.100
24	2.467	59	2.831	94	2.996	129	3.102
25	2.486	60	2.837	95	3.000	130	3.104
26	2.502	61	2.842	96	3.003	131	3.107
27	2.519	62	2.849	97	3.006	132	3.109
28	2.534	63	2.854	98	3.011	133	3.112
29	2.549	64	2.860	99	3.014	134	3.114
30	2.563	65	2.866	100	3.017	135	3.116
31	2.577	66	2.871	101	3.021	136	3.119
32	2.591	67	2.877	102	3.024	137	3.122
33	2.604	68	2.883	103	3.027	138	3.124
34	2.616	69	2.888	104	3.030	139	3.126
35	2.628	70	2.893	105	3.033	140	3.129
36	2.639	71	2.897	106	3.037	141	3.131
37	2.650	72	2.903	107	3.040	142	3.133
38	2.661	73	2.908	108	3.043	143	3.135
39	2.671	74	2.912	109	3.046	144	3.138
40	2.682	75	2.917	110	3.049	145	3.140
41	2.692	76	2.922	111	3.052	146	3.142
42	2.700	77	2.927	112	3.055	147	3.144
43	2.710	78	2.931	113	3.058	148	3.146
44	2.719	79	2.935	114	3.061	149	3.148

for specific details on how to account for zero flow years.

Confidence limits

A frequency curve is not an exact representation of the population curve. How well a stream record predicts flooding depends on record length, accuracy, and applicability of the underlying probability distribution. Statistical analysis allows the advantage of calculating confidence limits, which provide a measure of the uncertainty or spread in an estimate. These limits are a measure of the uncertainty of the discharge at a selected exceedance probability. For example, for the 5 percent and 95 percent confidence limit curves, there are nine chances in ten that the true value lies in the 90 percent confidence interval between the curves. As more data become available at a stream gage, the confidence limits will normally be narrowed. As presented in WRC Bulletin 17B, the following method is provided to develop confidence limits for a log-Pearson type III distribution.

$$X_{CI,U} = \bar{X}_l + S_l\left(K_{CI,U}\right) \qquad \text{(eq. 5–22)}$$

$$X_{CI,L} = \bar{X}_l + S_l\left(K_{CI,L}\right) \qquad \text{(eq. 5–23)}$$

where:

$X_{CI,U}$ = logarithmic upper confidence limit
$X_{CI,L}$ = logarithmic lower confidence limit
$\bar{X}_l$ = logarithmic peak flow mean
S_l = logarithmic peak flow standard deviation

$$K_{CI,U} = \frac{K_{LP,T} + \sqrt{K_{LP,T}^2 - ab}}{a} \qquad \text{(eq. 5–24)}$$

$$K_{CI,L} = \frac{K_{LP,T} - \sqrt{K_{LP,T}^2 - ab}}{a} \qquad \text{(eq. 5–25)}$$

$$a = 1 - \frac{z_c^2}{2\left(n-1\right)} \qquad \text{(eq. 5–26)}$$

$$b = K_{LP,T}^2 - \frac{z_c^2}{n} \qquad \text{(eq. 5–27)}$$

where:

n = record length
$K_{LP,T}$ = as listed in table 5–3, as a function of return period and skew coefficient
z_c = standard normal deviate, that is, the zero-skew $K_{LP,T}$ value at a return period of 1 decimal confidence limit. For the 95 percent confidence limit (0.05), z_c = 1.64485.

Comparisons of the frequency curve

Comparisons of distributions between the watershed being investigated and other regional watersheds can be useful for error checking and to identify possible violations of the underlying assumptions for the analysis. This can be especially illuminating for gages in the same watershed, upstream and downstream of the gage of interest, possibly identifying particular and unexpected hydrologic phenomena.

Discharge estimates from precipitation can be a helpful complement to gage data. However, such discharge estimates require a valid rainfall-runoff model. Such models are best when calibrated, which requires gage information. Such a calibrated model can be useful at other points within the watershed.

Example 4 illustrates analysis for outliers and confidence intervals.

Computation resources for flow frequency analysis

Several computer programs are available for assistance in performing the flood frequency analysis including the U.S. Army Corps of Engineers HEC–FFA (USACE 1992b) or the USGS PEAKFQ (USGS 1998). The USGS provides a computer program, PEAKFQ, at the Web site:

http://water.usgs.gov/software/peakfq.html

Spreadsheet programs have also been used to perform flood frequency analysis calculations as detailed by WRC Bulletin 17B. One of these spreadsheets is used in the examples presented in this chapter. This spreadsheet includes algorithms for generalized skew, 90 percent confidence intervals (95 percent confidence limits), historic data inclusion, and outlier identification. Example output sheets of this spreadsheet are provided in the example 5. It should be noted that these computational aids are for unregulated rivers and streams and that special precautions are necessary when evaluating flood frequencies on rivers with dams and significant diversions.

Transfer methods

Peak discharge frequency values are often needed at watershed locations other than the gaged location. Peak discharges may be extrapolated upstream or downstream from stream gages, for which frequency curves have been determined. In addition, peak dis-

Example 4: Confidence interval and outlier example

Problem: This example illustrates the analysis of USGS 06324500 Powder River at Moorhead, Montana, gage data in Eastern Montana (table 5–6). The following questions are addressed:

- frequency distribution for the gage
- 90 percent confidence interval
- outlier check
- impacts of the use of historic methodology and the impacts of inclusion of any outliers are assessed

Solution: The peak streamflow data were downloaded from the USGS NWIS data system (*http://waterdata.usgs.gov/nwis/*). These data, along with basic statistical computations, are provided in table 5–6.

Inspect the comments accompanying the peak flow data. For this data set, two of the data points are daily averages (instead of peak flows), six data points are estimates, and one estimate is an historic peak. The event on March, 17, 1979, is missing the peak flow (though the gage record does include the associated stage). Effort should be made to populate this point, but for this exercise, the data point is ignored.

It can be useful to plot frequency data to help in the identification of outliers and trends. Figure 5–2 includes a plot of these data.

This plot clearly shows a possible outlier and also qualitatively indicates a possible downward trend during the second half of the data set. A step trend would be more evident in the case of greater reservoir regulation within this watershed.

To assess the impact of using the historic methodology and inclusion or exclusion of the possible outlier, several frequency analyses need to be computed. A frequency analysis is performed on all of the data. Equation 5–9 is used to compute the distribution and confidence limits area, using equations 5–22 and 5–23. The computations and results are provided in table 5–6. The data are plotted in figure 5–3.

Outliers are identified using equations 5–5, 5–19, and 5–20, and the Weibull plotting positions are computed using equation 5–10. The high and low outlier thresholds are 10.96 and 6.48, respectively. Since the skew is between +0.4, high and low outliers are checked at the same time. The identification of outliers and computation of plotting position are shown in table 5–7. An outlier identified by the WRC Bulletin 17B methodology has been highlighted in yellow.

Since the 1923 event has been identified as a possible outlier, a frequency analysis is performed on a data set that excludes this high-flow value. The results of this computation are provided in table 5–8.

WRC Bulletin 17B provides a special methodology for historic peaks. The basis of this method is the assumption that data from the systematic record is representative of the period of the historic data. The systematic record and statistics are adjusted accordingly. This method is applied to the entire record of the Powder River at Moorhead gage, with computations that use equations 5–15 to 5–18 and those results are provided in table 5–9. Results have been plotted in figure 5–3.

Example 4: Confidence interval and outlier example—Continued

Inspection of the plotted results reveals a number of characteristics in the frequency distributions, specifically:

- The frequency analysis that includes the 1923 outlier in its computations (but does not incorporate the historic methodology) has the highest frequency distribution estimate. With the exception of the outlier, it also matches the higher data well. The 90 percent confidence interval brackets the higher data, with the exception of the outlier. This distribution does somewhat overestimate lower frequency events.
- The frequency analysis based on the historic methodology also well represents the higher data and somewhat overestimates lower data. This historic distribution is slightly lower than the nonhistoric distribution.
- The frequency analysis that excludes the outlier from its computations provides a distribution that is much lower than the distributions that include the outlier point. This distribution does not represent the higher flow data well—its 90 percent confidence interval excludes two additional data points, as plotted using the Weibull methodology. It does represent the lower peak data better.

It can be concluded from these observations that inclusion of the high outlier likely best represents the less frequent (higher) events. Exclusion of the data point provides a distribution that better represents more frequent (lower) events. For the Powder River at Moorhead, Montana, gaging station, it may be best to use the distribution that best represents the frequency of a desired event. If one distribution is required for all frequencies, the inclusion of the outlier using the historic methodology is likely best, due to its slightly better representation of all data than the nonhistoric, included outlier computation.

In addition, with the 1923 outlier perhaps being biased high, it may be best to revisit the computation of this historic peak. Additionally, it may be prudent to incorporate the generalized skew procedure to counteract any bias in the skew of the gage data.

Example 5: Log-Pearson spreadsheet frequency analysis example

The frequency analysis for the USGS gage 08251500 Rio Grande near Labatos, CO, is required for a stream stabilization project. The distribution was computed using a log-Pearson spreadsheet. The output sheets are provided in figures 5–4 to 5–6.

Visual observation of the graph of the plotted data indicates the computed record should be accepted for the analysis.

Table 5–6 Peak streamflow data at gage 06324500 Powder River at Moorhead, MT

Date	Peak flow (ft^3/s)	Notes	Date	Peak flow (ft^3/s)	Notes	Date	Peak flow (ft^3/s)	Notes
09/30/1923	100,000	2,3	06/15/1953	8,590		05/27/1980	2,210	
06/03/1929	8,610		08/06/1954	9,740		05/31/1981	2,160	
07/14/1930	4,040		06/18/1955	5,610		07/26/1982	6,350	
05/06/1931	6,040		06/16/1956	7,200		06/13/1983	2,870	
06/08/1932	3,550		06/07/1957	5,600		05/19/1984	4,620	
08/30/1933	14,800		06/12/1958	4,900		07/31/1985	1,410	
06/16/1934	1,920		03/19/1959	5,740		06/09/1986	4,540	
06/01/1935	8,140		03/20/1960	6,200		07/18/1987	11,400	
03/02/1936	9,240		05/30/1961	1,320		05/19/1988	1,990	
07/14/1937	14,500		06/17/1962	23,000		03/12/1989	800	1,2
05/30/1938	5,720		06/15/1963	7,010		08/21/1990	8,150	
06/02/1939	7,200		06/24/1964	15,000		06/04/1991	5,460	
06/04/1940	6,820		04/02/1965	18,300		11/12/1991	6,410	
08/13/1941	8,360		03/13/1966	4,000		06/09/1993	6,740	
06/26/1942	5,070		06/17/1967	17,300		07/09/1994	3,920	
03/26/1943	8,800		06/08/1968	8,580	2	05/11/1995	8,250	
05/20/1944	10,700		07/16/1969	5,280		03/13/1996	3,500	1,2
06/06/1945	6,190		05/24/1970	8,900	2	06/10/1997	4,290	
06/11/1946	5,720		06/01/1971	8,340		07/04/1998	2,760	
03/19/1947	9,300	2	02/29/1972	7,800		05/04/1999	3,960	
06/17/1948	9,320		06/19/1975	12,100		05/20/2000	3,930	
03/06/1949	9,360		06/23/1976	5,370		07/13/2001	1,490	
05/19/1950	2,620		05/17/1977	4,750				
09/09/1951	2,020		05/20/1978	33,000				
03/25/1952	15,300		03/17/1979					

Notes:
1/ Discharge is a maximum daily average
2/ Discharge is an estimate
3/ Discharge is an historic peak

Figure 5–2 Data plot at gage 06324500 Powder River at Moorhead, MT

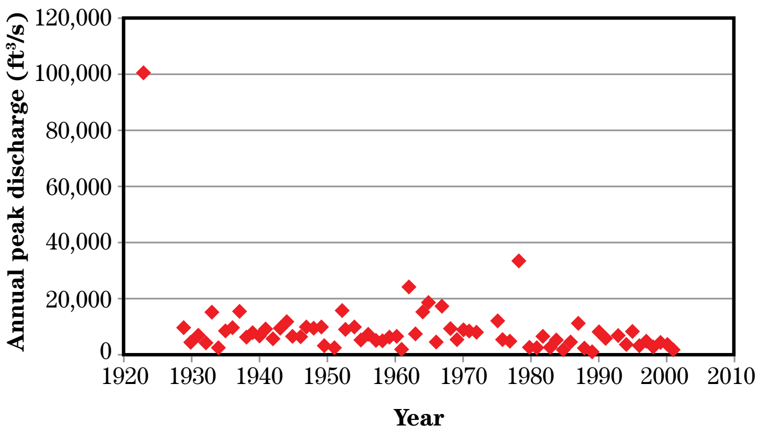


Figure 5–3 Data and frequency plots

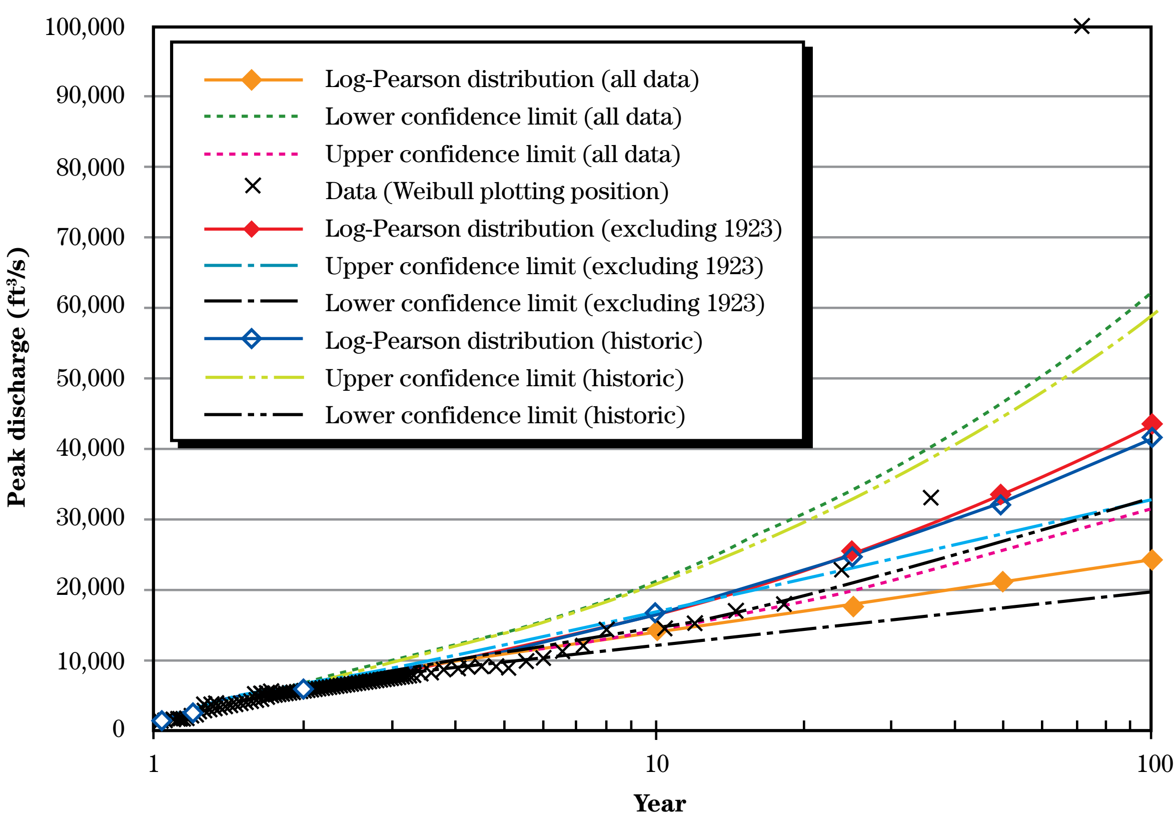


Table 5–7 Logarithmic data and Weibull plotting position values

Year	ln(Q)	Rank	Weibull (yr)
1923	11.513	1	72.0
1929	9.061	20	3.6
1930	8.304	53	1.4
1931	8.706	38	1.9
1932	8.175	58	1.2
1933	9.602	8	9.0
1934	7.560	67	1.1
1935	9.005	27	2.7
1936	9.131	17	4.2
1937	9.582	9	8.0
1938	8.652	40	1.8
1939	8.882	29	2.5
1940	8.828	32	2.3
1941	9.031	23	3.1
1942	8.531	47	1.5
1943	9.083	19	3.8
1944	9.278	12	6.0
1945	8.731	37	1.9
1946	8.652	41	1.8
1947	9.138	16	4.5
1948	9.140	15	4.8
1949	9.144	14	5.1
1950	7.871	62	1.2
1951	7.611	65	1.1
1952	9.636	6	12.0
1953	9.058	21	3.4
1954	9.184	13	5.5
1955	8.632	42	1.7
1956	8.882	30	2.4

Year	ln(Q)	Rank	Weibull (yr)
1957	8.631	43	1.7
1958	8.497	48	1.5
1959	8.655	39	1.8
1960	8.732	36	2.0
1961	7.185	70	1.0
1962	10.043	3	24.0
1963	8.855	31	2.3
1964	9.616	7	10.3
1965	9.815	4	18.0
1966	8.294	54	1.3
1967	9.758	5	14.4
1968	9.057	22	3.3
1969	8.572	46	1.6
1970	9.094	18	4.0
1971	9.029	24	3.0
1972	8.962	28	2.6
1975	9.401	10	7.2
1976	8.589	45	1.6
1977	8.466	49	1.5
1978	10.404	2	36.0
1980	7.701	63	1.1
1981	7.678	64	1.1
1982	8.756	35	2.1
1983	7.962	60	1.2
1984	8.438	50	1.4
1985	7.251	69	1.0
1986	8.421	51	1.4
1987	9.341	11	6.5
1988	7.596	66	1.1

Year	ln(Q)	Rank	Weibull (yr)
1989	6.685	71	1.0
1990	9.006	26	2.8
1991	8.605	44	1.6
1992	8.766	34	2.1
1993	8.816	33	2.2
1994	8.274	57	1.3
1995	9.018	25	2.9
1996	8.161	59	1.2
1997	8.364	52	1.4
1998	7.923	61	1.2
1999	8.284	55	1.3
2000	8.276	56	1.3
2001	7.307	68	1.1

Table 5–8 Frequency analysis data at gage 06324500 Powder River at Moorhead, MT

The basic statistics of the peak flow natural logarithms are provided:

Average =	8.7167	n =	71	
Standard deviation =	0.7722	z_c =	1.64485	(95% confidence limit)
Skew =	0.2883	a =	0.980674775	

Next, the log-Pearson K-Values are extracted from table 5–2, and the frequency values and confidence limits are computed:

Return period	**1.0526**	**1.25**	**2**	**5**	**10**	**25**	**50**	**100**	**200**
K-value	-1.559	-0.853	-0.048	0.825	1.308	1.845	2.205	2.536	2.845
ln (discharge)	7.513	8.058	8.680	9.354	9.727	10.142	10.419	10.675	10.914
Discharge (ft^3/s)	1,832	3,160	5,882	11,539	16,760	25,379	33,500	43,245	54,922
b	2.391	0.689	-0.036	0.642	1.673	3.367	4.824	6.391	8.057
$K_{CI,U}$	-1.293	-0.638	0.148	1.070	1.604	2.209	2.618	2.995	3.350
$K_{CI,L}$	-1.885	-1.101	-0.246	0.612	1.063	1.554	1.879	2.176	2.452
$\ln(Q_{CI,U})$	7.718	8.224	8.831	9.543	9.956	10.423	10.738	11.030	11.304
$\ln(Q_{CI,L})$	7.261	7.867	8.527	9.189	9.538	9.917	10.168	10.397	10.610
QCI,U (ft^3/s)	2,248	3,729	6,844	13,947	21,071	33,614	46,084	61,684	81,115
QCI,L (ft^3/s)	1,423	2,609	5,047	9,790	13,873	20,268	26,043	32,751	40,551

Table 5–9 Historic methodology computations

H = 79 Z = 1 n = 70 L = 0

From equation 5–15: systematic record weight, Ws = 1.114
From equation 5–16: historically adjusted average, $\tilde{X}$ = 8.713
From equation 5–17: historically adjusted standard deviation, $\tilde{S}$ = 0.765
From equation 5–18: historically adjusted skew, $\tilde{G}$ = 0.239
z_c =1.64485 (95% confidence limit) a = 0.980674775

Next, the log-Pearson K-Values are extracted from Table 5–5–2 and the frequency values and confidence limits are computed:

Return period	**1.0526**	**1.25**	**2**	**5**	**10**	**25**	**50**	**100**	**200**
K value	-1.573991189	-0.851162143	-0.039585477	0.827675714	1.304099048	1.830008811	2.179143811	2.499891431	2.799026432
ln (discharge)	7.509	8.062	8.682	9.346	9.710	10.112	10.379	10.624	10.853
Discharge (ft³/s)	1,824	3,171	5,898	11,448	16,480	24,639	32,180	41,126	51,697
b	2.439	0.686	-0.037	0.646	1.662	3.310	4.710	6.211	7.796
$K_{CI,U}$	-1.306	-0.636	0.158	1.075	1.601	2.193	2.589	2.955	3.298
KCI,L	-1.904	-1.100	-0.239	0.613	1.059	1.539	1.855	2.143	2.411
ln(QCI,U)	7.714	8.227	8.834	9.534	9.937	10.389	10.693	10.973	11.235
ln(QCI,L)	7.257	7.871	8.530	9.182	9.522	9.890	10.131	10.351	10.556
QCI,U (ft³/s)	2,239	3,739	6,861	13,827	20,682	32,516	44,036	58,261	75,707
QCI,L (ft³/s)	1,418	2,621	5,064	9,718	13,659	19,730	25,113	31,302	38,409

Figure 5–4 Sheet 1 of log-Pearson spreadsheet output for USGS gage 0825150

OUTPUT TABLES (The spreadsheet is configured so that only the area in these boxes will be printed.)

NRCS Log-Pearson Frequency Analysis Spreadsheet, Version 2.1, 5/2003 *Page 1 of 3*

Project: *Example 5-5-3*
Streamgage: # *USGS 08251500 RIO GRANDE NEAR LOBATOS, CO.*
Date: ######## **Performed By:** *Steve Yochum; Hydrologist, Northern Plains Engineering Team*

Without Generalized Skew

Average: 7.8440
Standard Deviation: 0.95782872
Skew Coefficient[1]: -0.3948897

Length of systematic record: 102
Number of historic peaks: 0
Length of Data Record: 102
Length of Historic Record:[5] ----

With Generalized Skew

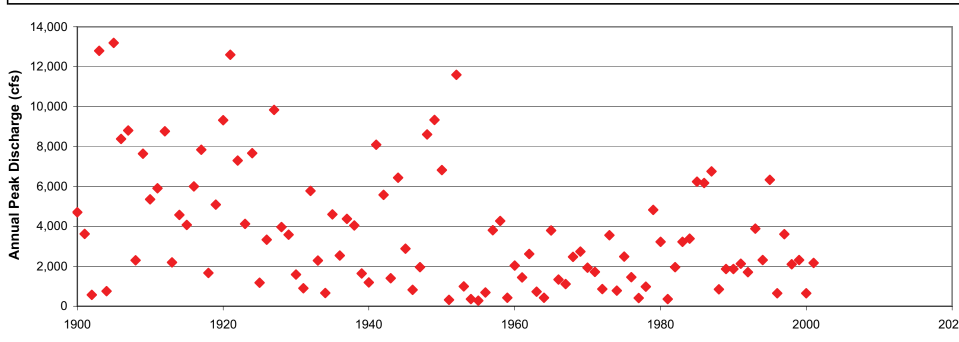
Generalized Skew Coefficient[3]: -0.0519
Variance of Generalized Skew[3]: 0.2408
A: -0.298409
B: 0.837329
station skew: -0.394890
MSE Station Skew: 0.07195506
Weighted skew coefficient[1]: -0.3159687

Recurrence Interval[2] (years)	Percent Chance	K-Value	Ln(Q)	Peak[4] Discharge (cfs)	90% Confidence Interval Upper (cfs)	90% Confidence Interval Lower (cfs)
200	0.5	2.206	9.9567	21,100	29,100	16,200
100	1	2.033	9.7911	17,900	24,200	13,900
50	2	1.837	9.6034	14,800	19,700	11,700
25	4	1.608	9.3841	11,900	15,400	9,560
10	10	1.232	9.0238	8,300	10,400	6,840
5	20	0.855	8.6628	5,780	7,030	4,870
2	50	0.065	7.9064	2,710	3,180	2,320
1.25	80	-0.816	7.0620	1,170	1,380	962
1.05	95	-1.749	6.1690	478	601	363
200	0.5	2.279	10.0270	22,600	31,500	17,300
100	1	2.092	9.8478	18,900	25,800	14,700
50	2	1.881	9.6457	15,500	20,600	12,200
25	4	1.637	9.4120	12,200	15,900	9,810
10	10	1.243	9.0343	8,390	10,500	6,910
5	20	0.853	8.6613	5,780	7,020	4,860
2	50	0.053	7.8943	2,680	3,140	2,290
1.25	80	-0.823	7.0560	1,160	1,380	956
1.05	95	-1.730	6.1871	486	611	370

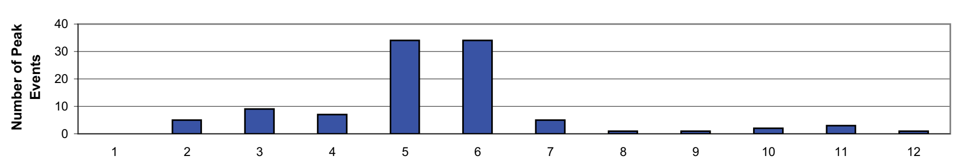
(1) Station and generalized skews must be between -2.00 and +3.00 in this spreadsheet.
(2) Considering the relatively short length of most gage records, less frequent peak estimates need to be used with considerable care.
(3) Computed one of four ways (see "generalized skew coefficient" worksheet): Mean and variance (standard deviation^2) of station skews coefficients in region; skew isolines drawn on a map or regions; skew prediction equations; read from Plate 1 of Bulletin 17B (reproduced in this spreadsheet), with MSE Generalized Skew = 0.302.
(4) Results are automatically rounded to three significant figures, the dominant number of significant figures in the K-Value table.
(5) Historic frequency analysis assu mes that intervening years reflect systematic record.

Comments: *(I) drainage area = 7700 mi^2, contributing drainage area (excluding closed basin) = 4760 mi^2.*
(ii) Generalized skew coefficient and variance computed by computing mean and variance of the station skews of the Upper Rio Grande basin.

Data Plot:



Peak Timing:



NRCS 2006

Figure 5–5 Sheet 2 of log-Pearson spreadsheet output for USGS gage 08251500

NRCS Log-Pearson Frequency Analysis Spreadsheet, Version 2.1, 5/2003 *Page 2 of 3*

Project: *Example 5-5-3*
Streamgage: # *USGS 08251500 RIO GRANDE NEAR LOBATOS, CO.*
Date: *5/27/2003* **Performed By:** *Steve Yochum; Hydrologist, Northern Plains Engineering Team*

Input Data **Station ID:** *08251500* **Latitude, Longitude:** *37°04'42"* *105°45'22"*
Drainage Area (mi^2): *4760* **County:** *Conejos*
Number of low outliers eliminated: *0* **State:** *CO*

	Date	Discharge (cfs)	Historic?	Outlier?		Date	Discharge (cfs)	Historic?	Outlier?		Date	Discharge (cfs)	Historic?	Outlier?
1	05/30/1900	4,700	n	n	51	03/30/1950	6,820	n	n	101	29560	650	n	n
2	05/23/1901	3,620	n	n	52	02/19/1951	320	n	n	102	30103	2170	n	n
3	05/15/1902	565	n	n	53	05/08/1952	11,600	n	n	103	----	----	n	n
4	06/18/1903	12,800	n	n	54	05/30/1953	995	n	n	104	----	----	n	n
5	04/19/1904	751	n	n	55	02/13/1954	360	n	n	105	----	----	n	n
6	06/08/1905	13,200	n	n	56	03/11/1955	280	n	n	106	----	----	n	n
7	06/17/1906	8,380	n	n	57	06/05/1956	681	n	n	107	----	----	n	n
8	07/03/1907	8,800	n	n	58	07/31/1957	3,810	n	n	108	----	----	n	n
9	06/14/1908	2,300	n	n	59	05/29/1958	4,270	n	n	109	----	----	n	n
10	06/10/1909	7,640	n	n	60	03/02/1959	418	n	n	110	----	----	n	n
11	04/30/1910	5,360	n	n	61	06/12/1960	2,040	n	n	111	----	----	n	n
12	06/13/1911	5,910	n	n	62	05/02/1961	1,440	n	n	112	----	----	n	n
13	05/29/1912	8,770	n	n	63	04/22/1962	2,620	n	n	113	----	----	n	n
14	03/23/1913	2,200	n	n	64	11/10/1962	724	n	n	114	----	----	n	n
15	06/05/1914	4,580	n	n	65	11/11/1963	423	n	n	115	----	----	n	n
16	05/19/1915	4,070	n	n	66	06/22/1965	3,790	n	n	116	----	----	n	n
17	05/12/1916	6,000	n	n	67	05/11/1966	1,330	n	n	117	----	----	n	n
18	06/20/1917	7,840	n	n	68	08/13/1967	1,110	n	n	118	----	----	n	n
19	06/16/1918	1,670	n	n	69	06/01/1968	2,470	n	n	119	----	----	n	n
20	05/25/1919	5,090	n	n	70	06/19/1969	2,730	n	n	120	----	----	n	n
21	05/27/1920	9,320	n	n	71	09/18/1970	1,930	n	n	121	----	----	n	n
22	06/16/1921	12,600	n	n	72	03/30/1971	1,720	n	n	122	----	----	n	n
23	06/01/1922	7,300	n	n	73	03/16/1972	856	n	n	123	----	----	n	n
24	06/17/1923	4,120	n	n	74	05/23/1973	3,560	n	n	124	----	----	n	n
25	05/21/1924	7,670	n	n	75	04/01/1974	784	n	n	125	----	----	n	n
26	02/14/1925	1,180	n	n	76	06/18/1975	2,490	n	n	126	----	----	n	n
27	06/04/1926	3,330	n	n	77	05/31/1976	1,450	n	n	127	----	----	n	n
28	07/03/1927	9,830	n	n	78	03/22/1977	405	n	n	128	----	----	n	n
29	06/01/1928	3,960	n	n	79	07/01/1978	979	n	n	129	----	----	n	n
30	05/27/1929	3,580	n	n	80	06/10/1979	4,830	n	n	130	----	----	n	n
31	06/01/1930	1,590	n	n	81	06/13/1980	3,230	n	n	131	----	----	n	n
32	03/22/1931	900	n	n	82	12/05/1980	360	n	n	132	----	----	n	n
33	05/24/1932	5,780	n	n	83	06/01/1982	1,950	n	n	133	----	----	n	n
34	06/03/1933	2,290	n	n	84	06/29/1983	3,230	n	n	134	----	----	n	n
35	02/19/1934	663	n	n	85	05/31/1984	3,390	n	n	135	----	----	n	n
36	06/18/1935	4,600	n	n	86	06/13/1985	6,240	n	n	136	----	----	n	n
37	05/07/1936	2,540	n	n	87	06/11/1986	6,180	n	n	137	----	----	n	n
38	05/19/1937	4,370	n	n	88	05/19/1987	6,760	n	n	138	----	----	n	n
39	05/02/1938	4,040	n	n	89	04/10/1988	848	n	n	139	----	----	n	n
40	03/24/1939	1,640	n	n	90	04/11/1989	1,870	n	n	140	----	----	n	n
41	05/19/1940	1,190	n	n	91	05/10/1990	1,860	n	n	141	----	----	n	n
42	05/16/1941	8,090	n	n	92	05/23/1991	2,130	n	n	142	----	----	n	n
43	05/13/1942	5,580	n	n	93	04/15/1992	1,700	n	n	143	----	----	n	n
44	05/04/1943	1,400	n	n	94	05/30/1993	3,890	n	n	144	----	----	n	n
45	05/18/1944	6,440	n	n	95	06/03/1994	2,320	n	n	145	----	----	n	n
46	05/12/1945	2,880	n	n	96	07/05/1995	6,330	n	n	146	----	----	n	n
47	11/12/1945	822	n	n	97	02/20/1996	650	n	n	147	----	----	n	n
48	05/11/1947	1,960	n	n	98	06/05/1997	3,610	n	n	148	----	----	n	n
49	06/07/1948	8,600	n	n	99	10/15/1997	2,100	n	n	149	----	----	n	n
50	06/22/1949	9,330	n	n	100	06/20/1999	2,310	n	n	150	----	----	n	n

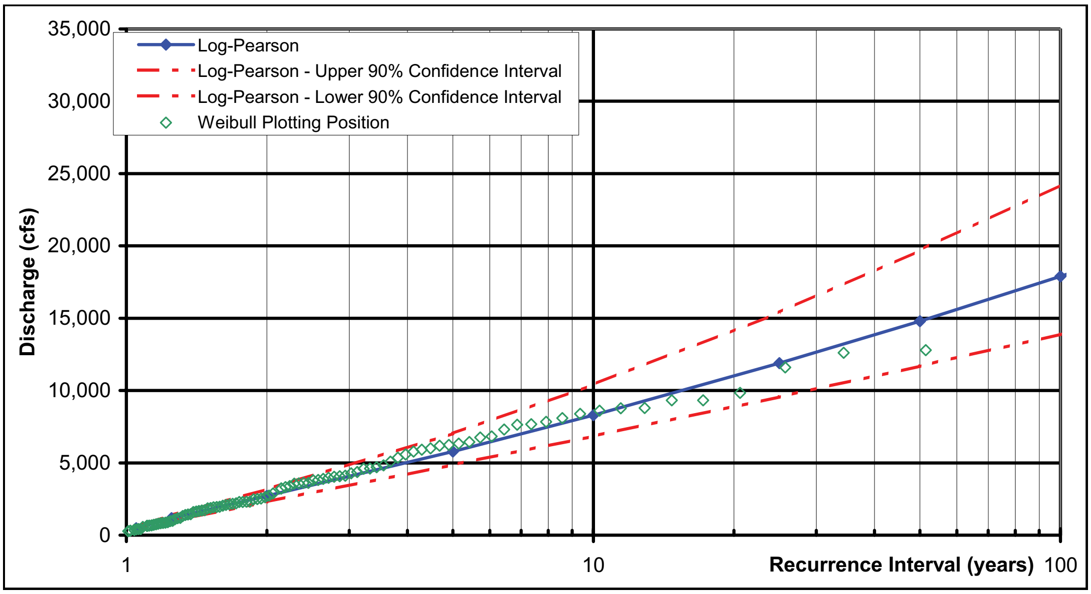
NRCS 2006

Figure 5–6 Sheet 3 of log-Pearson spreadsheet output for USGS gage 08251500

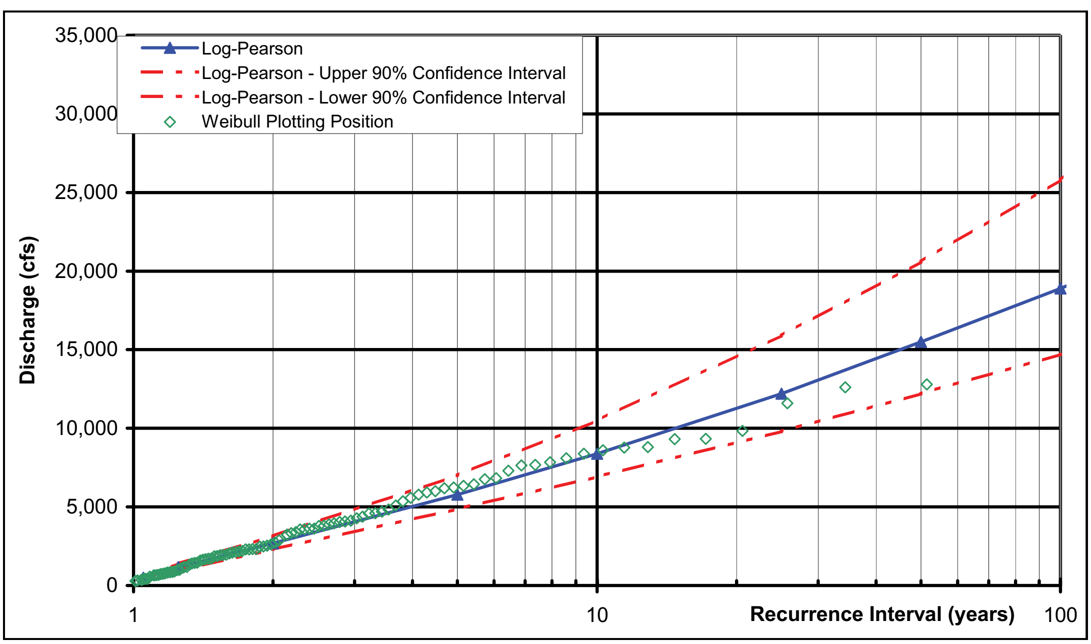
NRCS Log-Pearson Frequency Analysis Spreadsheet, Version 2.1, 5/2003 *Page 3 of 3*

Project: *Example 5-5-3*
Streamgage: # *USGS 08251500 RIO GRANDE NEAR LOBATOS, CO.*
Date: ######## **Performed By:** *Steve Yochum; Hydrologist, Northern Plains Engineering Team*

Discharge-Frequency, with Gage Skew
 # USGS 08251500 RIO GRANDE NEAR LOBATOS, CO



Discharge-Frequency, with Generalized Skew
 # USGS 08251500 RIO GRANDE NEAR LOBATOS, CO



NRCS 2006

charges may also be transferred or correlated from gage data from a nearby stream with similar basin characteristics.

Several equations and techniques exist for data transfer. Equation 5–28 is a simple transfer equation:

$$Q_u = Q_g \left(\frac{A_u}{A_g} \right)^x \qquad \text{(eq. 5–28)}$$

where:
Q_u = flood discharge at the ungaged stream
Q_g = flood discharge at the gaged stream
A_u = area at the ungaged stream
A_g = area at the gaged stream
x = regional exponent for area ratio (typically from 0.5 to 1)

Equation 5–28 can be used to develop comparative estimates. The regional exponent is computed by plotting a graph of flows for the same return period and similar basins, and then determining the slope of the best fit line on log–log paper. Example 6 illustrates the calculation. Again, specific regional data are needed for each state, and each hydrologic region within the state.

Transposition of peak flow rates is adversely affected by large differences in watershed lag times, runoff generated from small area thunderstorms, large differences in drainage area size, and differences in soils and vegetative cover. For transfer relations to be effective, the following conditions should be met:

- The drainage area ratio between the gaged and the ungaged area should be two or less.
- The watershed at the gaged location and ungaged watershed must be in the same climatic and physiographic region.

The more deviation exists from these two conditions, the more it is recommended that calculated values be compared with other sources such as regional regression data and computer models.

Low-flow frequency analysis

A project design may require a low-flow assessment for biological design elements or requirements. For example, it may be necessary to know the flow depths and velocities during a defined critical spawning time in a designed channel. A reference similar to WRC Bulletin 17B is not available for low-flow frequency analysis. However, the same log-Pearson type III frequency distribution, used for peak flow analysis, is often used for low-flow analysis.

In the United States, annual minimum flows usually occur in late summer and early fall. The annual minimum average flow for a specified number of consecutive days (usually 7 days) is the typical data point. Computationally, the annual minimum 7-day flow may be found as the annual minimum value of 7-day means. The USGS provides access to daily mean flow values at stream gages across the United States at:

http://waterdata.usgs.gov/nwis

As is done in peak discharge analysis, the mean and standard deviations of the logarithms of the data values are calculated. Then log-Pearson type III frequency factors are applied to assign frequencies or probabilities to the low-flow magnitudes. The frequencies are nonexceedance probabilities. Equation 5–1 is still applicable, but the practitioner needs to be cautioned regarding the meaning of the statistics. For example, the 50-year, 7-day low flow has a 2 percent annual chance of not being exceeded. For comparison, the 50-year peak flow has a 2 percent annual chance of being exceeded.

Inclusion of a period of substantial drought helps ensure that a data sample is representative for low-flow conditions. It should also be noted that the effects of basin development are relatively greater for low flows than for high flows.

Transfer of low-flow frequency estimates to ungaged sites is difficult because of the geologic influence on low flows. If basin characteristics are very similar, drainage area ratios may be used to transfer low-flow data from gaged to ungaged sites. A few low-flow measurements at the ungaged site are good verification for the transferred data.

Example 6: Transfer method

Problem: The gage site has a drainage area of 10 square miles (fig. 5–7). The designer needs an estimate of the 100-year flow at a drainage area of 20 square miles.

Solution: The regression equation developed from the regional data for 100-year flow values, developed from frequency analysis of stream gages in the area, is as follows:

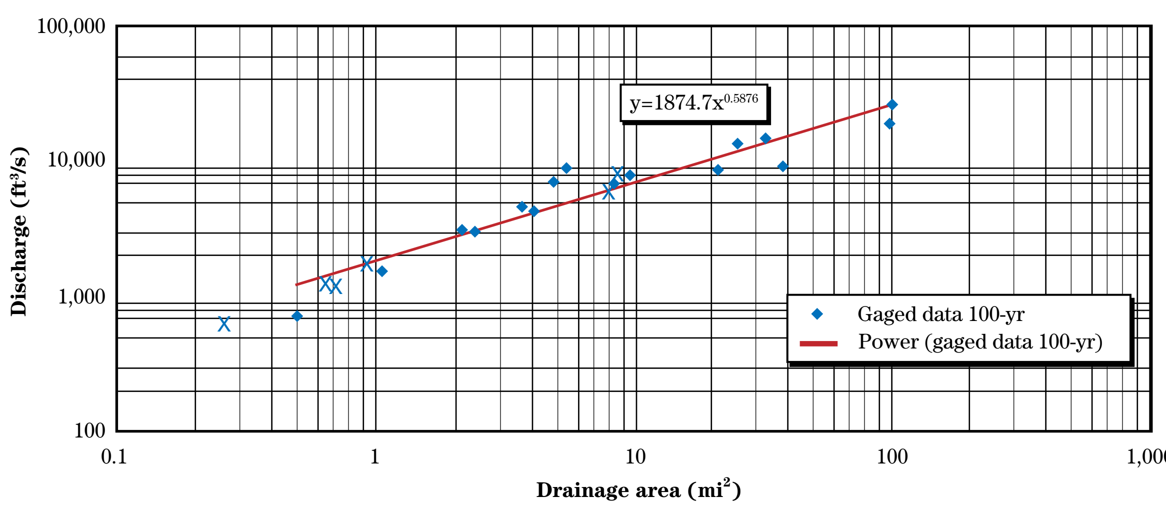
The value or intercept is the coefficient of the equation for a drainage area of 1 square mile (in log units = 0), and the intercept is 1,874.7.

The power of the equation is the slope of the line and is determined in log units as the discharge at 10 square miles, 7,253 cubic feet per second or 3.86 in log units. For a 1-square-mile drainage area, the discharge is 1,874.7 cubic feet per second and in log units is 3.27. The slope would be (3.86–3.27)/1 or 0.59.

Therefore, the 100-year flow at 20 square miles would be:

$$Q_u = 7,253\left(\frac{20}{10}\right)^{0.59} = 10,917 \text{ cubic feet per second}$$

Figure 5–7 100-yr discharges for the Rock Creek watershed in Montgomery County, MD



654.0505 Regional regression

Cost-effective designs for stream restoration, floodwater retarding structures, and many other conservation practices require peak streamflow frequency estimates. Peak streamflow frequency estimates represent peak discharges for return periods, generally ranging from 2 to 100 years. A regression equation for estimating peak discharges may be developed by statistically relating peak streamflow frequency and drainage basin characteristics for a geographic region of similar flood characteristics.

(a) Basic concepts

Regression forms
Regression is a method for developing a relationship between a dependent (Y) variable and one or more independent (X), predictor variables (NEH630.02, Hydrology). Regression assumptions are:

- No error exists in the independent variable; errors occur only in the dependent variable. Thus, regression is directional.
- Predictor variables are statistically independent.
- The observed values of the dependent variable are uncorrelated events.
- The population of the dependent variable is normally distributed about the regression line.
- A cause-and-effect relationship exists between predictor and dependent variables.

Regression is used to analyze hydrologic data because it provides an easy method for analyzing many factors simultaneously. The simplest form of the linear regression equation, with one predictor variable (X), is written as:

$$Y = a + bX \qquad \text{(eq. 5–29)}$$

where a and b are the intercept and slope regression coefficients. A more complicated form is the linear multiple regression equation, which relates a dependent variable and multiple predictor variables:

$$Y = b_0 + b_1X_1 + b_2X_2 \cdots + b_pX_p \qquad \text{(eq. 5–30)}$$

where:
Y = dependent variable, such as 100-year discharge
$b_0, b_1, b_2, \ldots, b_p$ = partial regression coefficients
$X_1, X_2, \ldots, X_p$ = independent (predictor) variables

Linear regression calculations are tedious by hand and are usually performed with the aid of programmed procedures on a computer. Example calculations may be found in NEH630.18.

Evaluating regression equations
One of the most commonly used measures of goodness of fit is the coefficient of determination, usually expressed as R^2. It is the dimensionless ratio of the explained variation in the dependent variable over the total variation of the dependent variable. A coefficient of determination of 1.0 indicates that the values of the dependent variable can be calculated exactly using the predictor variables in the given data set. The lower the R^2 value, the less direct the relation is and the wider the scatter in the data. Since this value is dimensionless, it can be used to compare goodness-of-fit of different regression equations. It does not provide a quantified expected variation. If a relationship is nonlinear, the regression coefficients will be dependent on the choice of independent variables, as well as on the curve fit relationship.

It should also be noted that a high degree of correlation (R^2 close to 1.0) does not necessarily mean that there is either causation or even a direct dependence between the variables. It only indicates that the given set of data can be predicted with the regression equation. In all circumstances, the reasonableness of the relationship between independent and dependent variables should be examined. Extremely high R^2 values (0.95 and above) can indicate bias in the data collection or an insufficient number of collected data points for the order of the calculated regression equation. For example, if only two data points are collected, a straight line regression equation between the two will have an R^2 value of 1.0.

Another measure of the quality of a regression equation is the standard error of estimate, typically expressed as $S_{Y,X}$. This is the root mean square of the estimates and is a measure of the scatter about the regression line of the independent variable. The standard error of estimate is not reflexive. It shows how well the dependent variable correlates to the independent variable, but not vice versa. The standard error of estimate has similar properties to the standard deviation and can be thought of as the standard deviation of the residuals. A residual is the difference between the value predicted with the regression equation and the observed dependent variable. As the standard error of estimate approaches 0, the quality of the regression equation increases.

Step-type regressions can be used to evaluate the significance of each predictor variable in a regression equation. The significance of adding or deleting predictor variables is evaluated with an F-test. A computed F greater than a table F-value indicates significance (see NEH 630.18 for more details). For example, a step forward regression starts with the most important predictor as the only variable in the equation. The most important of the remaining predictors is added, and the F-value computed. If this predictor is significant, another of the remaining predictors is added, and the process repeated. When a predictor is not found significant, the previous equation, not including that predictor, is used for analysis.

(b) Regional analysis

Regional study helps assure consistency of estimates at different locations and provides means for estimating discharge-frequency values at locations where gaged data are not available. Also, flow discharge estimates at a gaged location can usually be improved by a study of gaged frequency characteristics throughout the region.

Simplified regional study method
Regional analysis allows the estimation of peak discharge magnitude and frequency for ungaged watersheds by using relationships from nearby gaged watersheds. NEH630.02, Hydrology, provides the regional analysis in its simplest form.

- Select nearby gaged watersheds that are climatically and physically similar to the ungaged watershed.
- Construct frequency lines of peak discharges for each gaged watershed.
- Plot peak discharges for selected frequencies of each gaged watershed against its drainage area. Use log-log paper for plotting. A simple regression (curve fitting) between log of drainage area (predictor variable) and log of discharge (dependent variable) aids in drawing a best fit straight line for each selected frequency.
- Construct the frequency line for the ungaged watershed as follows: enter the plot with the ungaged drainage area, find and plot the discharges on log-probability paper, and draw the frequency line through the points.

Use of regression equations
Regression equations are used to transfer flood characteristics from gaged to ungaged sites through use of watershed and climatic characteristics as predictor variables. The USGS has developed regional regression equations for each state and some territories, usually as part of cooperative studies with state departments of transportation (USGS 2002a, Report 02–4168).

General descriptions of techniques that USGS uses in developing regression equations follow. Frequency lines of peak discharges are developed at gaging stations following the recommendations of WRC Bulletin 17B (WRC 1981).

The regression equations generally take the form:

$$Q_T = aX^bY^cZ^d \qquad \text{(eq. 5–31)}$$

where:
Q_T = peak discharge of selected frequency; 100-year discharge (dependent variable)
X,Y,Z = watershed or climatic characteristics (predictor variables)
a,b,c,d = regression coefficients

With the log transform, the equation takes the form of equation 5–30. The most often used watershed and climatic characteristics are drainage area, main channel slope, and mean annual precipitation. Regression regions are generally determined by using major watershed boundaries and an analysis of the areal distribution of the residuals. As noted above, residuals are the differences between regression and observed flow estimates. For USGS regression equations, the region has already been predetermined for the end user. Regression equation use is illustrated in example 7.

Example 7: USGS regional equation (USGS Report 96–4307)

An ungaged watershed is located in Region 5, Texas. The watershed has a drainage area (A) of 13.2 square miles and stream slope (SL) of 71.3 feet per mile, determined with the aid of USGS 7.5 min quadrangle maps. The following regression equation applies for estimating the 100-year discharge:

$$Q_{100} = 295A^{1.01}\left(SL\right)^{0.405} = 22{,}500 \text{ ft}^3/\text{s}$$

Report 96–4307 gives a standard error of estimate of 78 percent. This means that there is roughly a two-thirds chance that the true 100-year discharge falls between 4,950 cubic feet per second and 40,050 cubic feet per second. Report 96–4307 also gives a means to calculate more exact confidence intervals (not shown here). An output report for the same ungaged watershed, generated by the USGS National Flood Frequency Program, follows:

```
National Flood Frequency Program
Version 3.0
Based on Water-Resources Investigations Report 02-4168
Equations from database NFFv3.mdb
Updated by kries 10/16/2002 at 3:51:06 PM new equation from WRIR 02-4140
Equations for Texas developed using English units

Site: MilldamTX, Texas
User: lgoertz
Date: Thursday, July 10, 2003 04:26 PM

Rural Estimate: Rural 1
  Basin Drainage Area: 13.2 mi2
  1 Region
  Region: Region_5_(A<32mi2_(51km2))
   Contributing_Drainage_Area = 13.2 mi2
   Stream_Slope = 71.3 ft/mi

Flood Peak Discharges, in cubic feet per second

             Recurrence     Peak,  Standard   Equivalent
Estimate     Interval, yrs  cfs    Error, %   Years
___________ _____________  _____  ________   __________

Rural 1            2         919      75
                   5        2910      63
                  10        5180      66
                  25        9310      69
                  50       15900      72
                 100       22500      78
                 500       50600
```

Accuracy and limitations
The standard errors of estimate or prediction range from 30 to 60 percent for most regression equations. The largest standard errors generally are for equations developed for the western part of the Nation, where the variability of the flood records is greater, gaging stations are less dense, and flood records are generally shorter. Regression equations developed from gaged natural basins should only be used on natural basins to make regression estimates. A natural basin may be defined as a basin with less than 10 percent impervious cover and less than 10 percent of its drainage area controlled or manipulated to affect peak stream flow. Users should exercise caution in extrapolating flood estimates beyond the ranges of predictor variables used in developing the equations.

Regression equations are not as accurate as frequency analysis from gaged data. For design purposes in high risk situations, both regression equations and hydrologic modeling methods should be employed.

(c) Computational resources for regional regression analysis of peak flows

The USGS has developed and published regression equations for a variety of locations within the United States. These equations have been compiled into the National Flood Frequency (NFF) Program. A computer program, National Flood Frequency Program, version 3: A Computer Program for Estimating Magnitude and Frequency of Floods for Ungaged Sites provides access to this information. The following USGS Web site provides regional regressions for flood peaks developed for many regions throughout the United States:

http://water.usgs.gov/software/nff.html

Regional regression relationships for bankfull discharge
Bankfull discharge regional regression relationships can present some different issues to a designer than relationships for peak flows. However, the basics remain the same. Information on developing regional regression relationships for bankfull discharge is provided in NEH654 TS5.

654.0506 Flow duration

Flow duration is the percentage of time that a given flow was equaled or exceeded over a period of time. A flow-duration curve for stream flow represents the hydrograph of the average year (or season) with its flows arranged in order of magnitude. For example, the flow value in the average year to be exceeded 20 percent of the time may be read from the flow-duration curve for that location.

Flow-duration curves have been used in the analysis of sediment transport quantities, critical habitat functions, water quality management alternatives, and water availability. It is often important to determine how the proposed restoration project will perform with low or normal flows. While flow-duration curves are typically calculated for several (usually >10) years of homogeneous record, they can be developed for specific seasons since seasonal flow variations can have critical habitat importance. For example, a project goal may include a minimum flow depth during a critical spawning period for anadromous fish species and a lower minimum depth for resident fish species. The same techniques used to develop flow-duration curves for sediment analysis can also be used to assess and design for habitat conditions. An example is provided in figure 5–8.

The USGS has developed flow-duration curves for many gaged locations in the United States. These curves are normally available on request from the USGS. The construction procedure used by USGS is outlined in Searcy (1959). Procedures for developing flow-duration curves are also described in Hydrologic Frequency Analysis, EM 1110–2–1415 (USACE 1993a). Data are typically sorted by magnitude, and the percent of the time that each value is exceeded is calculated. Since the data points are typically daily averages, each point will not necessarily be independent. This is a relatively simple statistical analysis. Data bin ranges are developed, and the numbers of occurrences are counted for each bin. For example, the number of times the flow was between 500 and 1,000 cubic feet per second would be counted. Then the percent of occurrence is assembled as a cumulative distribution function to define the percent of time that the flow is above a certain discharge or level. Flow-duration analysis is performed by using daily average flow (or other periods such as 3-day, 5-day, or weekly) during the period of interest. Historical and real time daily average flow data can be found at the following USGS Daily Flow Stream Gage Data Web site:

http://nwis.waterdata.usgs.gov/usa/nwis/discharge

The data can be stratified by seasons for this analysis, depending on study goals. The information can be for the entire period of record. However, if the watershed has undergone some significant change (as is typical in many stream restoration projects), it may be necessary to use only the record since the change has occurred. This is necessary to keep the data homogeneous.

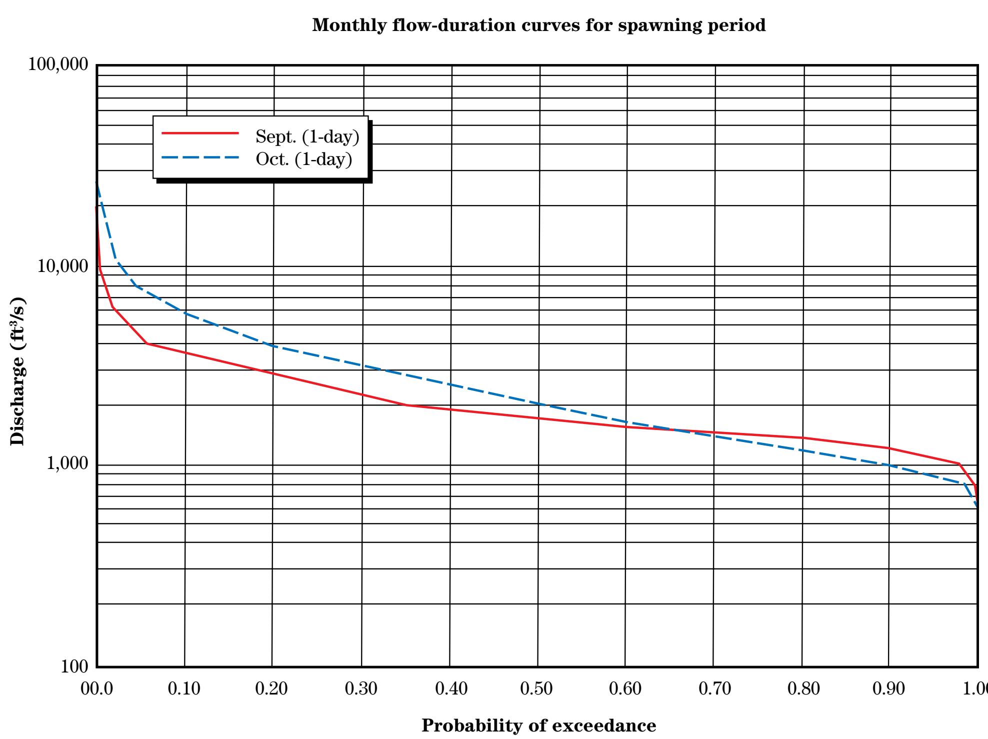
Transfer methods, as described earlier, can be used to transfer flow duration information from gaged sites to ungaged areas. However, these should have similar watershed characteristics, and the ratio of gaged to ungaged drainage area should be between 0.5 and 2.0 for reliable results. The accuracy of such a procedure is directly related to the similarity of the two sites. Typically, there is more error in transferring or estimating the ends of a flow-duration curve. Flow duration is dependent on watershed conditions. If regional flow-duration relations are to be developed, it is recommended that a measure of watershed conditions be included as an independent variable.

Two methods for estimating a flow-duration curve for ungaged sites are described by Biedenharn et al. (2000). They are the:

- drainage area flow-duration curve method
- regionalized-duration curve method

Graphs for the drainage area flow-duration curve method, for a specified recurrence interval discharge versus drainage area, are developed for a number of sites on the same stream or within hydrological similar portions of the same drainage basin. If data are reason-

Figure 5–8 Typical flow-duration curve



ably homogeneous, regression techniques should be used to generate curves of flow for selected percentile versus drainage area. By knowing the drainage area of the selected site(s), a flow-duration curve can be generated from the regression equations.

With the regionalized duration curve method, a nondimensional flow-duration curve is developed for a hydrologically similar gaged site by dividing discharge by bankfull discharge or by a specified recurrence interval discharge. Then a specified recurrence interval discharge is computed for the ungaged site using the aforementioned regression equations. Finally, the flow-duration curve for the ungaged site is derived by multiplying the dimensionless flows (Q/Q_2) from the nondimensional curve by the site Q_2. It should be noted that both methods simply provide an approximation to the true flow-duration curve for the site because perfect hydrologic similarity never occurs.

654.0507 Hydrologic models

There are many mathematical and computer hydrologic modeling systems available for predicting runoff from precipitation and snowmelt events that provide the volume and timing of water moving through the system. Models provide the ability to estimate existing, as well as future rainfall runoff patterns for a variety of conditions. Depending on the hydrologic model used, either single event peak flow or continuous multiple event modeling can be performed.

The accuracy of models is highly dependent on calibration data, which can often be difficult to acquire. However, if the issues that are to be addressed are comparative in nature rather than absolute, the importance of calibration is diminished. However, the results of a model study should fall between the USGS regional regression equation for the site and the upper bounds of one standard error of estimate. If the results of the model calibration are not within these bounds, after adjustment of the model parameters within reasonable limits, the reasons for the final answer and its derivation must be explained in the project documentation.

The level of accuracy required for a specific hydrologic analysis generally depends on the specific characteristics of each individual project. The selection of the appropriate methodology should be done with a firm understanding of the assumptions, accuracy, data requirements, and limitations of the approach. Brief statements on the use of the models are provided.

The rational method (rational formula) is one of the easiest models to implement. It can be used for drainage areas up to 80 hectares (200 acres). Use of the rational formula on larger drainage areas requires sound judgment to ensure reasonable results. The hydrologic assumptions underlying the rational formula include:

- constant and uniform rainfall over the entire basin
- a rainfall duration equal to the time of concentration

The rational method is not appropriate if:

- the basin has more than one main drainage channel

- the basin is divided so that hydrologic properties are significantly different in one section versus another
- the time of concentration is greater than 60 minutes
- storage is an important factor

The NRCS TR–55 method (USDA Soil Conservation Service (SCS) 1986) provides a manual method for computing peak discharges for drainage basins. The TR–55 method is segmental (flow time is computed by adding the travel times for the overland, shallow concentrated, and channel segments). TR–55 considers hydrologic parameters such as slope of the watershed and channel, channel roughness, water losses, rainfall intensity, soil type, land use, and time. TR–55 should be used with caution when the design is highly sensitive to the computed peak flow values. TR–55 also assumes that rainfall is uniform over the entire basin. Additional assumptions include:

- the basin is drained by a single main channel or by multiple channels with times of concentration that are nearly equal
- the weighted curve number should be greater than 40
- runoff from snowmelt or rain on frozen ground cannot be estimated using the procedures in TR–55
- the time of concentration should be between 0.1 and 10 hours
- storage in the drainage area is less than 5 percent of the runoff volume and does not affect the time of concentration
- a single composite curve number can accurately represent the watershed runoff characteristics

A computer program has been developed to automate the manual procedures in TR–55. The computer program developed in the Windows® environment is known as WinTR–55. The WinTR–55 computer program is available at the following Web site:

http://www.wcc.nrcs.usda.gov/hydro/hydro-tools-models-wintr55.html

The HEC–1/HEC–HMS models are rainfall-runoff models developed by the USACE Hydrologic Engineering Center (USACE 1981). These models can be used with basins of almost any size and complexity. HEC–1 is designed to simulate the surface runoff resulting from precipitation over a watershed by representing that watershed as an interconnected system of components. These components consist of surface runoff, stream channels, and reservoirs. Each component is represented by a set of parameters, which specify its characteristics, and the mathematical relations, which describe its physical processes. The end result of the HEC–1 modeling process is the computation of runoff hydrographs for the subbasins and stream channels. The program is composed of five basic sub models as illustrated in figure 5–9.

HEC–1 assumes that the rainfall is spatially uniform over each subbasin modeled. NRCS rainfall time distributions, loss methods, dimensionless unit hydrographs, and the lag equations often are used; however, careful consideration must be given to the assumptions and limitations underlying these methods. For example, the NRCS has published an upper limit on basin size for the NRCS lag equation of 800 hectares (2,000 acres, 3.1 mi^2) (NEH630.15). The upper limit on basin area for the NRCS Loss Method (runoff curve number) is not well established; however, a limit of 20 square miles has been suggested. These limitations may be overcome by subdivision of the watershed and appropriate routing. Various GIS packages can be used as an interface to HEC–1. These GIS techniques systematize the computation of the physiographic and hydrologic parameters required by HEC–1.

The WinTR–20 model is a rainfall-runoff model developed by the NRCS (USDA NRCS 2004). It can be used with basins of almost any size and complexity. WinTR–20 is designed to simulate the surface runoff resulting from precipitation over a watershed by representing that watershed as an interconnected system of components. These components consist of surface runoff stream channels and reservoirs. The program is composed of five submodels as illustrated in figure 5–9. Normally, it is assumed that the rainfall is uniform over each subbasin. However, that rainfall total can be varied for each subbasin. Actual or design temporal rainfall distributions can be used with standard dimensionless unit hydrographs that are part of the normal inputs. Several GIS computer programs that can be used to develop the areal input values such as curve numbers are available. NRCS has developed a GIS computer program that provides geographical information in the proper format for WinTR–20 (USDA NRCS 2004). The program is available from the following Web site:

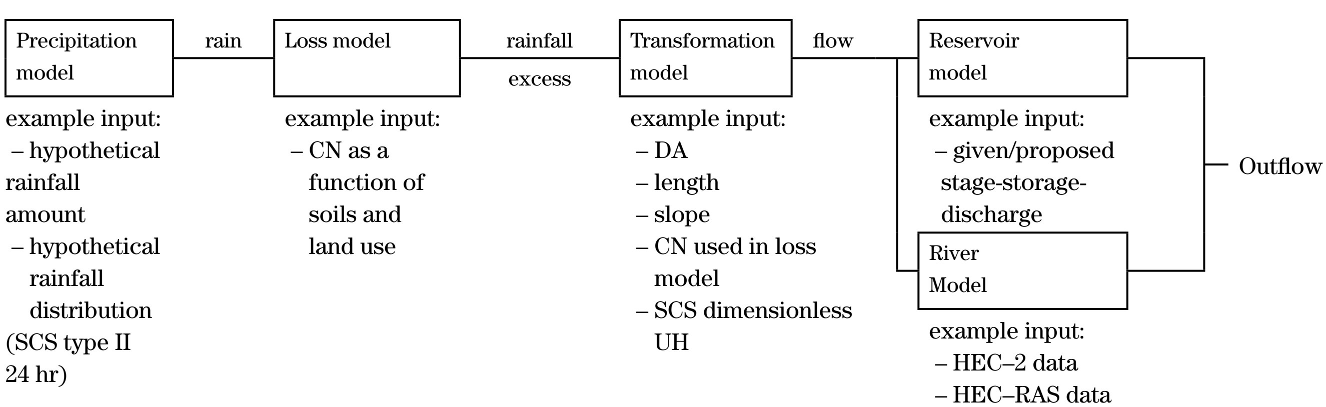
www.wcc.nrcs.usda.gov/hydro/

Use of these models is fairly common in ungaged systems, or in areas where land use and stormwater detention systems significantly alter the hydrograph. With the advent and collection of soil, vegetation, topography and land use types in GIS, model development and database management is a simpler process.

654.0508 Channel-forming discharge

Natural alluvial streams experience a wide range of discharges and adjust their shape and size during flow events that have sufficient energy to mobilize the channel boundary materials. Until the 1960s, it was widely assumed that floods of great magnitude, but low frequency, controlled channel form because of the nonlinear relationship between discharge and sediment transport capacity. Sediment transport increases exponentially with discharge. This view was challenged by Wolman and Miller (1960) who argued that in most streams, over an extended period of time, the total amount of sediment transported by a discharge of a given magnitude depends not only on its transport capacity, but also its frequency of occurrence. Thus, although extremely large events can produce spectacularly high sediment loads, they happen so infrequently and last such a short time that their overall contribution to the total sediment movement during a long period is relatively small. Small events also make a small contribution to the total sediment moved because their high frequency of occurrence is offset by their very low sediment transport capacity. It follows from this logic that flow of both moderate magnitude and moderate frequency is responsible for the greatest

Figure 5–9 Five basic submodels of a rainfall/runoff model



amount of sediment movement (Leopold, Wolman, and Miller 1964). However, recent studies have indicated that this concept may not hold true for all streams (Werrity 1997).

Channel-forming discharge concept

The channel-forming discharge concept is based on the idea that, for a given alluvial channel, there exists a single steady discharge that, given enough time, would produce channel dimensions equivalent to those produced by the natural hydrograph. This discharge is thought to dominate channel form and process. Estimates of channel-forming discharges are used to classify stream types, estimate channel dimensions, assess stability, and express hydraulic geometry relationships.

While many techniques and methodologies are used to estimate a channel-forming discharge in stable alluvial channels, all can be characterized as one of four main types. These are:

- discharge based on bankfull indices
- discharge based on drainage area
- discharge based on specified statistical recurrence intervals
- discharge based on an effective discharge calculation

Discharge based on bankfull indices

Channel-forming discharge based on bankfull indices is determined by visually inspecting the reach in question or surveys of this reach to locate morphological evidence of the bankfull stage. The discharge associated with this stage is then computed or estimated. Identifying relevant features that define the bankfull stage can be problematic (Williams 1978), particularly in dynamic, unstable channels (Simon, Dickerson, and Heins 2004). Many field indicators have been proposed and are briefly described in table 5–10.

Identifying bankfull stage from indicators is subjective. None of the bankfull indicators is applicable in all situations (Williams 1978). Many workers use a combination of the indices in an iterative fashion. However, even experienced observers may arrive at conflicting or misleading results, particularly for conditions outlined in table 5–11.

The field identification of bankfull indicators is particularly problematic in stream reaches that are unstable or threshold. If the project reach is not stable or alluvial, it may be possible to find indicators of bankfull stage in stable alluvial reaches upstream or downstream. However, since stream restoration is most often practiced in unstable watersheds, field determination of bankfull stage may be impractical or impossible (Copeland et al. 2001). An exception could be found in a stable and alluvial incised stream that has formed a new flood plain within the incised channel. In this case, the top of the high bank is now an abandoned flood plain or terrace, and there should be newly formed top-of-bank features within the older incised channel. However, it is important to remember that the new flood plain may not yet be fully formed; that is, the channel may not be stable and may still be aggrading. In addition, a new inset flood plain (sometimes referred to as incipient flood plain) may be restricted in width or height due to channel constraints. Measurements taken in such situations would give misleading values for the bankfull discharge.

When applying the estimate of bankfull stage from one reach to another, it is important to keep in mind that the location of the break between the channel and the flood plain is influenced by many factors, including (but not limited to) the following:

- climatic regime (humid vs. arid)
- geologic erosion conditions of the streambank materials (bedrock vs. unconsolidated material; coarse vs. fine textures; cohesive vs. noncohesive)
- stream slope
- hydrologic regime (perennial vs. intermittent versus ephemeral)
- sediment source, quantity and supply including distribution along the active channel and flood plain.
- stream confinement or width
- stream downcutting or incisement
- size and type of vegetation on the flood plain and within the channel
- controls on channel width and alignment such as riprap and bridge abutments

Table 5–10 Summary of bankfull indices

Bankfull indicator	Reference
Minimum width-to-depth ratio	Wolman (1955) Pickup and Warner (1976)
Highest elevation of channel bars	Wolman and Leopold (1957)
Elevation of middle bench in rivers with several overflow sections	Woodyer (1968)
Minimum width-to-depth ratio plus a discontinuity (vegetative and or physical) in the channel boundary	Wolman (1955)
Elevation of upper limit of sand-sized particles in boundary sediment	Leopold and Skibitzke (1967)
Elevation of low bench	Schumm (1960); Bray (1972)
Elevation of active flood plain	Wolman and Leopold (1957) Nixon (1959)
Lower limit of perennial vegetation	Schumm (1960)
Change in vegetation (herbs, grass, shrubs)	Leopold (1994)
A combination of: • elevation associated with the highest depositional features • break in bank slope • change in bank material • small benches and other inundation features • staining on rocks • exposed root hairs	Rosgen (1994)

Table 5–11 Summary of stream conditions that affect bankfull indices

Reach condition	Process	Effect on bankfull indices
Threshold	Sediment transport capacity of the reach exceeds the sediment supply, but the channel grade is stable	Bankfull indices may be relics of extreme flood events, and may indicate a bankfull flow that is too high
Degrading	The sediment transport capacity of the reach exceeds the sediment supply to the reach, and the channel grade is lowering	The former flood plain is in the process of becoming a terrace. As a result, bankfull indices may indicate a flow that is too high
Aggrading	The sediment transport capacity of the reach is less than the sediment supply	The existing flood plain or in channel deposits may indicate a flow that is too low
Recently experienced a large flow event	Erosion and/or deposition may have occurred on the bed and banks	Bankfull indices may be missing or may reflect the large flow event
Channelized	Sediment transport capacity may not be in balance with sediment supply. The channel may be aggrading or degrading. The reach may be functioning as a threshold channel	Bankfull indices may be relics of previous channel, artifacts of the construction effort, embryonic, or missing altogether

- controls on channel depth and slope such as drop structures, rock weirs, check dams, beaver dams, and cross vanes.

For example, the bankfull discharge measured from a reach with a narrow flood plain may be inappropriate for use on another reach of the same stream, which has a wide flood plain.

Once bankfull stages are estimated for a stream reach (generally over at least one meander wavelength or 10 channel widths); the bankfull discharge can be estimated. This is often done with either a resistance formula calculation such as Manning's equation or with a computer model such as HEC–RAS (Brunner 2002). Practitioners should keep in mind that the use of resistance equations such as Manning's equation or the Darcy-Weisbach equation, while rapid, are subject to the error inherent to the normal depth assumption. In addition, it should be noted that because stage is not a unique function of discharge in alluvial streams, some data scatter should be expected (Copeland et al. 2001). Uncertainty associated with stage-discharge relationships is addressed in more detail in standard manuals and texts (USACE 1996). Additional guidance on the identification of bankfull discharge indicators is provided in NEH654 TS5.

Discharge based on drainage area

Many relationships are available that correlate dominant discharge to drainage area. These offer a quick technique for assessing a dominant discharge. However, the practitioner should keep in mind that these relationships are basically best fit lines that are plotted through a data set. There is a distribution of valid bankfull discharge estimates that will fall both above and below the line. For example, figure 5–10 illustrates such a curve developed by Emmett (1975) for the Salmon River in Idaho. Although the regression line fits the data in a visually satisfactory fashion, it should be noted that for a drainage area of about 70 square miles, the bankfull discharge varied between about 300 and 900 cubic feet per second.

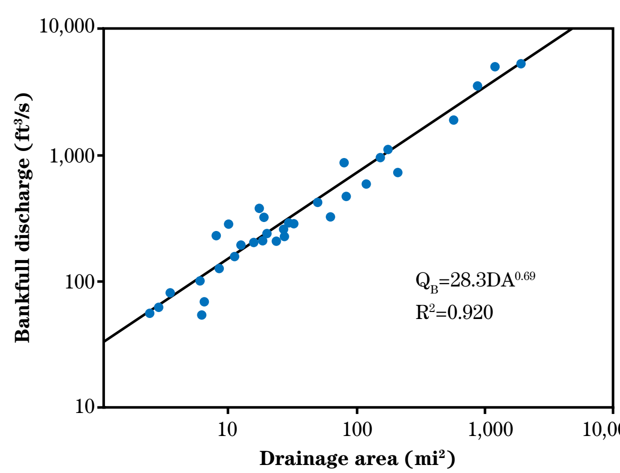
While drainage area is certainly an important factor in estimating streamflow, it is only one of many parameters affecting runoff. Caution should also be used when assessing the relevance of the relationship to watersheds in different physiographic areas or watersheds with different runoff characteristics. Finally, while drainage area is certainly an important factor to estimate streamflow, it is only one of many parameters affecting runoff.

Discharge based on a specific recurrence interval

Many practitioners have related the channel-forming discharge to a specific recurrence interval. The use of a recurrence interval to estimate the channel-forming discharge offers the advantage of being able to calculate a value using gage records, hydrologic modeling, or regional regression relations. Regression equations for estimating discharges with recurrence intervals from 2 to 100 years (Q_2 to Q_{100}) are available for the entire United States via *http://water.usgs.gov/software/nff.html*, as well as from many state and local organizations. This recurrence interval for channel-forming discharge is often assumed to correspond to fall between Q_1 and $Q_{2.5}$, with a mean of $Q_{1.5}$ (Leopold 1994).

However, there are many instances where the channel-forming discharge does not fall within the 1- to 2.5-year range. Williams (1978) showed that out of 35 flood plains he studied in the United States, the bankfull discharge (measured at top of bank) varied between the 1.01- and 32-year recurrence interval, and that only about a third of those streams had a bankfull discharge recurrence interval between 1 and 5 years.

Figure 5–10 Bankfull discharge as a function of drainage area for the Salmon River, ID



In a similar study, Pickup and Warner (1976) showed that bankfull recurrence intervals ranged from 4 to 10 years. The recurrence interval is usually calculated by determining the flow that corresponds to bankfull indices as addressed in the previous section. Therefore, the issues addressed that are associated with the reliable physical identification of bankfull discharge indices impact the calculation of the recurrence interval and may account for some of the discrepancies. Simon, Dickerson, and Heins (2004) used computations based on suspended-sediment transport to compute effective discharge for 10 gages on unstable sand-bed channels in Mississippi. The resulting values of effective discharge ranged from 0.56 to 2.72 of the $Q_{1.5}$, with a mean of 1.04 $Q_{1.5}$.

Nevertheless, the use of a specified recurrence interval is often used as a first approximation of channel-forming discharge. But, because of the noted discrepancies, field verification is generally recommended to ensure that the selected discharge reflects morphologically significant features.

Discharge based on an effective discharge calculation

The effective discharge is defined as the mean of the arithmetic discharge increment that transports the largest fraction of the annual sediment load over a period of years (Andrews 1980). The effective discharge incorporates the principle prescribed by Wolman and Miller (1960) that the channel-forming discharge is a function of both the magnitude of the event and its frequency of occurrence. An advantage of using the effective discharge is that it is a calculated value not subject to the problems associated with determining field indicators (Copeland et al. 2001). Effective discharge computation consists of three steps.

Step 1 The flow-duration curve is derived from available stream gage data.

Step 2 Sediment data or an appropriate sediment-transport function is used to construct a bed material sediment rating curve.

Step 3 The flow-duration curve and the bed material sediment rating curve are integrated to produce a sediment load histogram that displays sediment load as a function of discharge for the period of record. The histogram peak is the effective discharge increment.

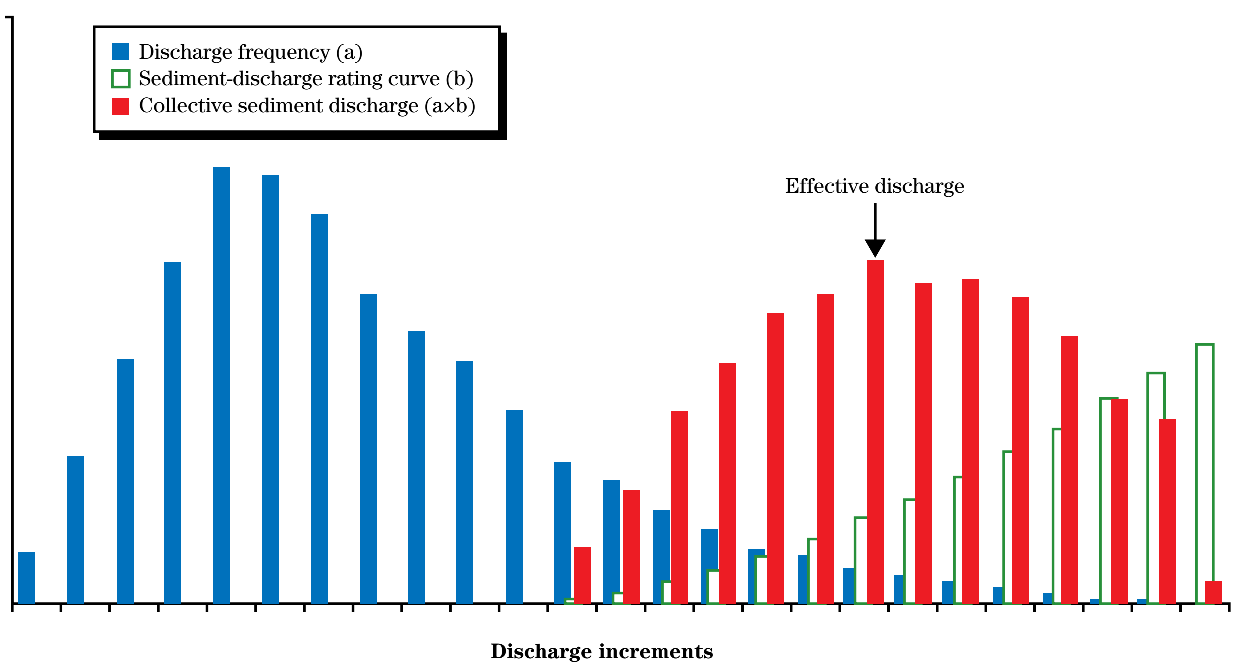
Specific instructions for calculating effective discharge can be found in the literature (Copeland et al. 2001; Biedenharn et al. 2000; and Thomas et al. 2000). Details of the procedure can influence the outcome, so study of these references is recommended. A graphical representation of the relationship between sediment transport, frequency of the transport, and the effective discharge is shown in figure 5–11. The peak of the effective discharge curve in figure 5–11 marks the discharge fraction that transports most of the material, and therefore, does the most work in forming the channel.

Effective discharge analyses may be performed for ungaged reaches by synthesizing a flow-duration curve and applying an appropriate sediment transport function to obtain a bed material sediment rating curve. Flow-duration curve synthesis may be done by plotting curves of discharge versus upstream drainage area for a given exceedance duration, using data from gages within the same watershed as the site of interest. A family of such curves may be created by varying the exceedance duration, and an appropriate flow-duration curve for the site of interest may be interpolated using its drainage area (Hey 1975). If flow-duration data are not available for adjacent gages, then regional information may be used after dividing discharge by either bankfull discharge or the 2-year discharge to produce a dimensionless ratio (Watson, Dubler, and Abt 1997). The dimensionless curve may be applied to the site of interest by multiplying by the base (Q_2 or bankfull Q) that is estimated using one of the aforementioned methods.

Since channel instability is the result of an imbalance in sediment supply and transport capacity, the greatest advantage of using effective discharge in restoration design lies in the fact that it requires quantification of the sediment transport capacity of a channel for a given hydrologic regime. Various channel geometries can be examined for their competence to transport the incoming sediment load, facilitating comparison of permutations of channel dimensions to optimize sediment transport efficiency within logistical constraints. This information is also useful when predicting the impact of alteration of watershed conditions with respect to sediment loads (upstream dam removal) or hydrology (urbanization) on channel stability (Copeland et al. 2001).

An important limitation of using an effective discharge analysis is that it is based on the assumption that the stream will transport the amount of sediment that it is hydraulically capable of moving, and it is this hydraulic capacity that forms the channel. In an urbanized watershed, once the urbanization is complete, the result is that the drainage area is partially covered so that the overland sediment yield reduced. In the Piedmont Region of Maryland, for example, many streams have degraded to bedrock or contain bed material that has been winnowed to a coarse gravel or cobble. These conditions, coupled with an increase in average annual flows, indicate that streams may have an excess sediment transport capacity. In this situation, the channel may be now operating as a threshold channel and the concept of effective discharge may not be relevant. Additional errors occur in effective discharge computations due to the assumption that sediment discharge is a continuous function of water discharge. Internal fluvial system thresholds or limitations on sediment supply may invalidate this assumption, leading to major errors at higher discharges (Nash 1994). An example calculation of effective discharge is provided in example 8.

Figure 5–11 Effective discharge calculation



Example 8: Effective discharge

Problem: Given the following flow-duration curve (fig. 5–12) and sediment transport rating curve (fig. 5–13), calculate the effective discharge:

Solution: The sediment transport rating curve was calculated from data collected during field surveys. The bed material gradation in the upstream supply reach was determined from the average of three volumetric bulk samples taken laterally across the stream. The cross-sectional geometry and slope were surveyed. Hydraulic parameters were calculated assuming normal depth. The Meyer-Peter Muller equation was chosen to make the sediment calculations because the bed was primarily gravel. The calculated bed material sediment transport rating curve is shown in figure 5–13.

The basic approach is to divide the natural range of streamflows during the period of record into a number of arithmetic classes, and then calculate the total bed material quantity transported by each class. This is achieved by multiplying the frequency of occurrence of each flow class by the median sediment load for that flow class. This can be accomplished using a spreadsheet or the USACE SAM program (Thomas, Copeland, and McComas 2003).

Table 5–12 represents output from the SAM program for the given conditions. The discharge increment with the largest increment of sediment transport is between 1,000 and 1,200 cubic feet per second. The effective discharge is then 1,000 cubic feet per second. The program also calculates the average annual sediment load, which is the sum of the sediment loads for each increment. In this case, the annual sediment load is 10,677 tons. A graphical representation of the effective discharge calculation is shown in figure 5–14.

Figure 5–13 Sediment transport rating curve calculated from bed material gradation collected upstream from the project reach and hydraulic parameters from surveyed cross section

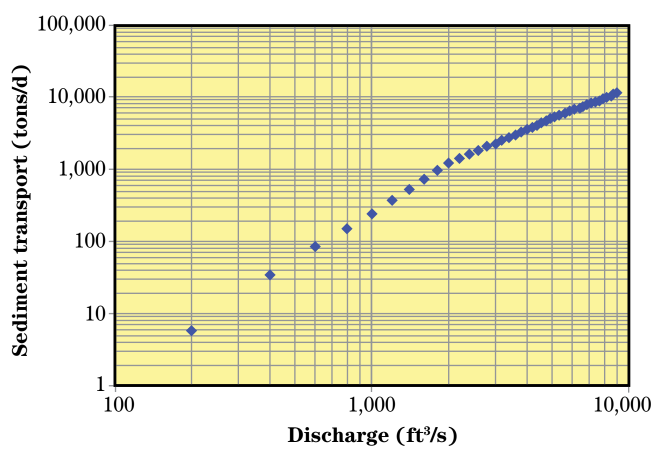


Figure 5–12 Flow-duration curve developed from 39 years of record at a USGS gage downstream from the project reach

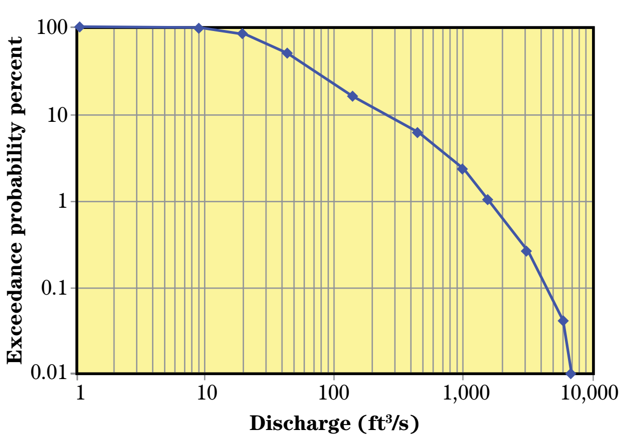


Figure 5–14 Effective discharge calculation

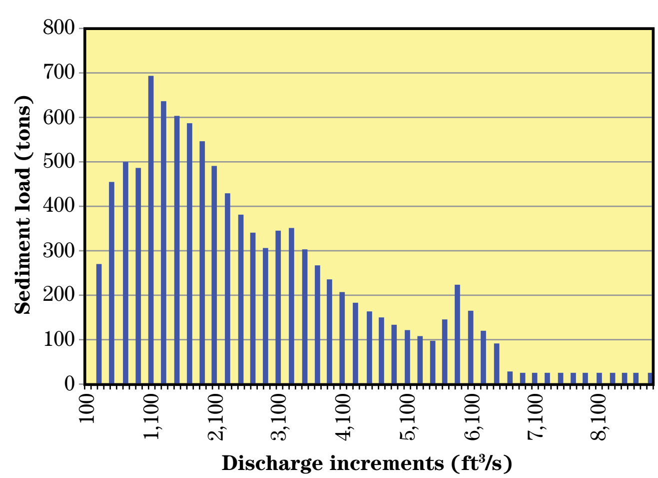


Table 5–12 Effective discharge calculation from SAM program

```
****************************************************************************************************
*                     SAMwin Software Registered to the US Army Corps of Engineers                 *
****************************************************************************************************
*                                    SEDIMENT YIELD CALCULATIONS                                   *
*                                            Version 1.0                                           *
*                       A Product of the Flood Control Channels Research Program                   *
*             Coastal & Hydraulics Laboratory, USAE Engineer Research & Development Center         *
*                                        in cooperation with                                       *
*                            Owen Ayres & Associates, Inc., Ft. Collins, CO                        *
****************************************************************************************************
TABLE 2.1 SEDIMENT DISCHARGE TABLE.
          Q,CFS    =           10.0     20.0     50.0     100.0     200.0
     QS,TONS/DAY   =   0.0      0.0        0.0               0.0       2.3

          Q,CFS    =           500.0    1100.0            2000.0    5000.0    10000.0
     QS,TONS/DAY   =   42.7    283.1            1074.6   4428.0    11178.0
TABLE 2.2 FLOW-DURATION TABLE
    #     CFS      %        #    CFS       %        #     CFS        %

    1       0.00    97.10  5      137.00     15.90     9    3090.00     0.25
    2      20.10    84.10  6      442.00      6.00    10    9000.00     0.00
    3      22.00    50.00  7      988.00      2.30
    4      44.20    50.00  8     1545.00      1.00

TABLE 2.3 INTEGRATION PARAMETERS FOR FLOW-DURATION OPTION
                         MINIMUM FLOW, CFS    =          0.00
                         MAXIMUM FLOW, CFS    =       9000.00
                 INTEGRATION INTERVAL, CFS    =         24.66
               NUMBER OF INTEGRATION STEPS    =           365
TABLE 2.7 DENSITY OF SEDIMENT DEPOSIT.
      IN LB/CUFT         =          93.00
      IN CY/TON          =           0.80
     TABLE 3.1 CALCULATED YIELDS
     SEDIMENT TRANSPORT FUNCTION USED -- MPM(1948),D50
    TIME PERIOD,           DAYS     =   354.415
    WATER YIELD,           ACFT     =   84445.,     Mean Daily Flow,     CFS     =  120.13
    SEDIMENT YIELD,        TONS     =   10677.,      Mean Daily Load,    T/D     =  30.
                           CUYD     =      8504.,    Mean Daily Conc,   mg/l     =  92.880
TABLE 3.2 DISTRIBUTION OF YIELD BY WATER DISCHARGE CLASS INTERVAL.
     NO. OF CLASSES    =       45      ,       CLASS INTERVAL =       200.00
     MINIMUM Q, CFS    =       0.00,   MAXIMUM Q,     CFS     =       9000.00
```

Table 5–12 Effective discharge calculation from SAM program—Continued

Class	Discharge ft^3/s	Sediment tons/d	Increment of water		Increment of sediment		
			acre-ft	%	%	tons	yd^3
	0	0					
1			12	0.01	0	0	0
	200	2					
2			14263	16.89	1.2	128	102
	400	21					
3			11422	13.53	3.0	320	255
	600	66					
4			8311	9.84	3.87	414	329
	800	132					
5			6288	7.45	4.19	448	356
	1000	225					
6			7033	8.33	6.18	660	526
	1200	344					
7			5136	6.08	5.6	598	476
	1400	485					
8			3993	4.73	5.2	555	442
	1600	653					
9			3293	3.9	5.0	534	425
	1800	849					
10			2633	3.12	4.59	490	390
	2000	1075					
11			2154	2.55	4.11	439	349
	2200	1245					
12			1795	2.13	3.6	384	306
	2400	1424					
13			1518	1.8	3.19	340	271
	2600	1612					
14			1301	1.54	2.85	304	242
	2800	1807					
15			1128	1.34	2.57	274	218
	3000	2011					
16			719	0.85	1.7	181	144
	3200	2222					
17			464	0.55	1.13	121	96
	3400	2440					
18			464	0.55	1.17	125	99
	3600	2665					
19			464	0.55	1.21	129	103

Table 5–12 Effective discharge calculation from SAM program—Continued

Class	Discharge ft^3/s	Sediment tons/d	Increment of water		Increment of sediment		
			acre-ft	%	%	tons	yd^3
	3800	2897					
20			464	0.55	1.24	132	105
	4000	3137					
21			464	0.55	1.27	136	108
	4200	3382					
22			464	0.55	1.31	140	111
	4400	3634					
23			464	0.55	1.34	143	114
	4600	3893					
24			464	0.55	1.37	147	117
	4800	4157					
25			464	0.55	1.41	150	119
	5000	4428					
26			464	0.55	1.43	153	122
	5200	4666					
27			464	0.55	1.45	155	123
	5400	4907					
28			464	0.55	1.47	157	125
	5600	5152					
29			464	0.55	1.48	159	126
	5800	5399					
30			464	0.55	1.5	160	128
	6000	5649					
31			464	0.55	1.52	162	129
	6200	5902					
32			464	0.55	1.54	164	131
	6400	6158					
33			464	0.55	1.55	166	132
	6600	6416					
34			464	0.55	1.57	167	133
	6800	6677					
35			464	0.55	1.58	169	135
	7000	6941					
36			464	0.55	1.6	171	136
	7200	7207					
37			464	0.55	1.61	172	137
	7400	7476					
38			464	0.55	1.63	174	138
	7600	7747					
39			464	0.55	1.64	175	140
	7800	8021					
40			464	0.55	1.66	177	141

Table 5–12 Effective discharge calculation from SAM program—Continued

Class	Discharge ft^3/s	Sediment tons/d	Increment of water		Increment of sediment		
			acre-ft	%	%	tons	yd^3
	8000	8297					
41			464	0.55	1.67	178	142
	8200	8575					
42			464	0.55	1.68	180	143
	8400	8855					
43			464	0.55	1.7	181	144
	8600	9138					
44			464	0.55	1.71	183	146
	8800	9423					
45			464	0.55	1.72	184	147
	9000	9710					
Total			84445	100	100	10677	8504

(a) Cautions and limitations

In addition to the previously noted limitations associated with the methods of estimation, other precautions should be applied to the entire nature of channel-forming discharge. The channel-forming discharge concept is based on the idea that there exists a single steady discharge that, given enough time, would produce channel dimensions equivalent to those produced by the natural long-term hydrograph. Although conceptually attractive, this definition is not necessarily physically feasible because riparian vegetation, bank stability, and even the bed configuration would be different in a natural stream than in a stream with a constant discharge (Copeland et al. 2001).

In addition, it is important to note that extreme events often have the capability to move a significant amount of sediment and cause major changes in channel cross section, profile, and planform. In streams that have experienced catastrophic events, the flow-frequency and sediment-transport relations may have changed or be changing with time as the channel adjusts. Results obtained using any technique may represent a condition that does not accurately depict present flow and sediment-transport conditions.

To design a stream restoration project with long-term stability, it is necessary to evaluate the full range of flows that will affect the channel. Therefore, stable channel design includes the evaluation of sediment transport capacity for a range of flows (not just the design discharge) to determine whether the project will aggrade or degrade, as well as meet other objectives for the restoration.

654.0509 Other sources of design flows

Other sources may include estimates of local flows. These can range from hydrologic models conducted as part of another study to historical records of extreme events. Regulatory or legislatively defined flows may be defined. Data may be available for irrigation releases, dam operations or navigation controls. It may be necessary to include an analysis of some or all of these flows. However, it is important to review this information to assess both the technical accuracy, as well as the assumptions made in their estimate. Additional calculations are often required.

654.0510 Conclusion

Rarely does the behavior of a channel under a single discharge adequately reflect the range of design conditions required for a stream restoration project. The design capacity of the channel should consider environmental objectives, as well as flood criteria. Often, habitat features are designed to narrow the channel during summer low flows to increase habitat during a biologically critical period. Project features are designed to withstand a significant flood event, normally a 10-year frequency discharge or larger. A realigned channel is normally designed to convey a flow selected for channel stability, normally larger than the 1-year frequency discharge. If the stream channel is realigned or reconstructed, a suitable design discharge must be selected. In many situations, this is the channel-forming discharge. A wide variety of sources and techniques exist for obtaining hydrologic data available to the designer. If a gage is available and the conditions applicable, a gage analysis is generally considered preferable, since it represents actual data for the stream. However, it is important to assess the applicability of the historic gage data to the current project conditions. For example, rapid increases of imperviousness in an urban watershed may have increased flows and resulted in stream instability. If this is the case, the historic gage data must not be used, because there is no realistic way to adjust the peak flow frequency to predevelopment conditions. Changes in rainfall-runoff characteristics may render historic gage data obsolete. Gage records provide an actual representation of the hydrologic behavior of a watershed. However, when a gage record is of short duration, or poor quality, or the results are judged to be inconsistent with field observations or sound judgment, then the analysis of the gage record should be supplemented with other methods.

Several state and local agencies have developed regional regression relations to estimate peak discharges at ungaged sites. These data can be readily applied, but care must be taken to assure that the regression relations include relevant parameters that can relate the unique characteristics of the study watershed to the data that were used to create the relations. Care must also be taken to make sure that the watershed parameters of the ungaged watershed being analyzed are within the watershed parameters used to develop the regression curves. It is also important to assess the relevance of the confidence limits of the estimates to the project analysis.

Hydrologic models provide the ability to estimate existing, as well as future rainfall runoff patterns for a variety of conditions. The use of models is preferred in cases where the watershed has changed. The accuracy of models is dependent on calibration data, which can be difficult to acquire. However, if the issues to be addressed are comparative in nature rather than absolute, the importance of calibration is diminished. The level of accuracy required for a specific hydrologic analysis generally depends on the specific characteristics of each individual project. The appropriate methodology should be selected with a firm understanding of the assumptions, accuracy, data requirements, and limitations of the approach. The designer should consult with a hydraulic engineer before deciding on which procedure should be used to obtain the needed flow data.

Channel-forming discharge can be estimated using a prescribed methodology. All methodologies for estimating channel-forming discharge present challenges. The practitioner should review the assumptions, data requirements and consider his or her experience when determining which technique to use. It is recommended that all available methods be used and cross checked against each other to reduce the uncertainty in the final estimate of the channel-forming flow.